

STELLAR KINEMATICS OF $z \sim 2$ GALAXIES AND THE INSIDE-OUT GROWTH OF QUIESCENT GALAXIES^{1,2}JESSE VAN DE SANDE³, MARISKA KRIEK⁴, MARIJN FRANX³, PIETER G. VAN DOKKUM⁵, RACHEL BEZANSON⁵, RYCHARD J. BOUWENS³, RYAN F. QUADRI⁶,
HANS-WALTER RIX⁷, ROSALIND E. SKELTON⁵*Submitted to ApJ*

ABSTRACT

Using stellar kinematics measurements, we investigate whether massive, quiescent galaxies were denser at $z \sim 2$ than they are today. We present X-Shooter spectra from the UV to NIR and dynamical mass measurements of 5 quiescent massive ($> 10^{11} M_{\odot}$) galaxies at $z \sim 2$. This triples the sample of $z > 1.5$ galaxies with well constrained ($\delta\sigma < 100 \text{ km s}^{-1}$) velocity dispersion measurements. From spectral population synthesis modeling we find that these galaxies have stellar ages that range from 0.5-2 Gyr, with no sign of on-going star formation. We measure velocity dispersions ($290\text{--}450 \text{ km s}^{-1}$) and find that they are 1.6-2.1 times higher than those of galaxies in the SDSS at fixed mass. Sizes are measured using GALFIT from HST-WFC3 H_{160} and UDS K-band images. The dynamical masses correspond well to the SED-based stellar masses, with dynamical masses that are $\sim 15\%$ higher. We find that M_*/M_{dyn} may decrease slightly with time, which could reflect the increase of the dark matter fraction within an effective radius. We combine different stellar kinematic studies from the literature, and examine the structural evolution from $z \sim 2$ to $z \sim 0$: we confirm that at fixed dynamical mass, the effective radius increases by a factor of ~ 2.8 , and the velocity dispersion decreases by a factor of ~ 1.7 with time. The mass density within one effective radius decreases by a factor of ~ 21 , while within a fixed physical radius (1 kpc) it decreases only mildly (factor of ~ 2.3). When we allow for an evolving mass limit by selecting a population of galaxies at fixed number density, a stronger size growth with time is found (factor of ~ 4), velocity dispersion decreases by a factor of ~ 1.4 , and interestingly, the mass density within 1 kpc is consistent with no evolution. This finding suggests that massive quiescent galaxies at $z \sim 2$ grow in an inside-out matter, consistent with the expectations from minor mergers.

Subject headings: cosmology: observations — galaxies: evolution — galaxies: formation

1. INTRODUCTION

Recent studies have shown that a considerable fraction of massive galaxies at $1.5 < z < 2.5$ have quiescent stellar populations (e.g. Labbé et al. 2005; Kriek et al. 2006; Williams et al. 2009). Among the most massive galaxies ($M_* > 10^{11} M_{\odot}$) approximately 40% are no longer forming stars (e.g. Whitaker et al. 2011; Brammer et al. 2011). Surprisingly, these massive quiescent galaxies have been found to be extremely compact (e.g. Daddi et al. 2005; Trujillo et al. 2006; van Dokkum et al. 2008; Franx et al. 2008; van der Wel et al. 2008; and numerous others), compared to their likely present-day counterparts.

Searches for any ultra-dense low-redshift counterparts by Trujillo et al. (2009) and Taylor et al. (2010a) found only a handful of compact sources at $z \sim 0$, that have relatively young stellar populations (Trujillo et al. 2009; Ferré-Mateu et al. 2012). The dearth of massive, old compact objects at low redshift implies that massive galaxies must have undergone severe structural evolution in size.

Errors in the size estimates have been invoked as a possible explanation for the compactness of massive high-redshift galaxies. Initial concerns that the size may have been underestimated, due to an envelope of low surface brightness light, have been addressed with deep HST WFC3 imaging (Szomoru et al. 2010; 2012), and by stacking results (e.g. van der Wel et al. 2008; Cassata et al. 2010; van Dokkum et al. 2008, 2010). The light could also be more concentrated due to the presence of active galactic nuclei (AGN) in these galaxies. However, spectra of these galaxies have shown that the light is dominated by evolved stellar populations, not AGN (Kriek et al. 2006, 2009; van de Sande et al. 2011; Onodera et al. 2012).

The question of whether stellar masses are accurate out to $z \sim 2$ remains, however, of paramount importance: an overestimate in stellar mass would bring the galaxies closer to the $z \sim 0$ mass-size relation. To date, basically all (stellar) masses have been derived by fitting the spectral energy distributions. This method suffers from many systematic uncertainties in stellar population synthesis models (SPS) (e.g. Conroy et al. 2009; Muzzin et al. 2009) and is essentially untested at $z > 1.5$.

Direct stellar kinematic mass measurements, which do not suffer from these uncertainties, can be derived by measuring the galaxy's velocity dispersion and the shape and extent of its luminosity profile, i.e. the Sérsic n parameter and effective radius. In particular, for low-redshift galaxies in the SDSS, Taylor et al. (2010b) showed that stellar mass is a very good predictor of dynamical mass, but only when non-homology of luminosity profile is properly accounted for if we include a Sérsic dependent virial factor (e.g. Cappellari et al. 2006). Although dynamical measurements of massive galaxies are

¹ Based on X-Shooter-VLT observations collected at the European Southern Observatory, Paranal, Chile.

² Based on observations with the NASA/ESA Hubble Space Telescope (HST), obtained at the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS 5-26555.

³ Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands.

⁴ Astronomy Department, Berkeley, Hearst Field Annex, Berkeley, CA 94720-3411, USA

⁵ Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520-8101, USA.

⁶ Carnegie Observatories, Pasadena, CA 91101, USA

⁷ Max Planck Institute for Astronomy, Königstuhl 17, D- 69117 Heidelberg, Germany

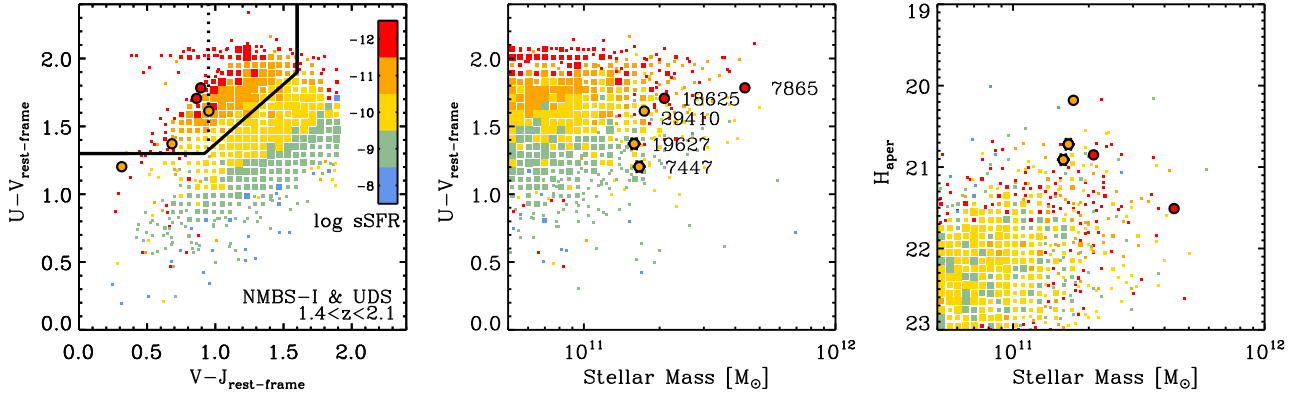


FIG. 1.— Comparison of our spectroscopic sample to the full population at similar redshift. Squares represent the density of galaxies from the NMBS-I and UDS at $1.4 < z < 2.1$ with mass $> 10^{10.5} M_{\odot}$. *Left panel:* Rest-frame U-V and V-J colors. Color coding is based upon the sSFR derived from SED fitting, red colors indicates low sSFR (quiescent), and blue colors high sSFR (star-forming). Galaxies in the top left region, as marked by the black line, all have low sSFR rates. This region is therefore often used to select quiescent galaxies at high redshift (Williams et al. 2009). All but one of our galaxies fall within this region, but their sSFR indicate that they are all truly quiescent. The vertical dotted line discriminates between young post-starburst like (left) versus old quiescent (right) as indicated by Whitaker et al. (2012). The strong Balmer absorption lines spectroscopically confirm the young ages of this sample. *Middle panel:* Rest-frame U-V versus stellar mass. At fixed mass, we find that most of our galaxies have similar colors to the entire population, except for NMBS-COS7447 and UDS-19627 on the blue side. *Right panel:* H -band aperture magnitude versus stellar mass. It is clear that our sample was selected on magnitude, and at fixed mass they are truly among the brightest galaxies, consistent with their post-starburst nature.

TABLE 1
TARGETS AND OBSERVATIONS

Catalog	ID	α	δ	Exp. Time [min]	J_{aper}	H_{aper}	K_{aper}	K_{tot}	$U - V_{\text{rf}}$	$V - J_{\text{rf}}$	$24 \mu\text{m}$ [μJy]	$\text{SFR}_{24\mu\text{m}}$ $M_{\odot} \text{ yr}^{-1}$
NMBS-COS	7447	10:00:06.96	02:17:33.77	120	21.09	20.72	20.63	19.64	1.20	0.31	$\lesssim 18$	$\lesssim 13$
NMBS-COS	18265	10:00:40.83	02:28:52.15	90	22.67	20.85	20.61	19.62	1.71	0.86	$\lesssim 18$	$\lesssim 15$
NMBS-COS	7865	10:00:17.73	02:17:52.75	434	22.75	21.51	21.07	20.02	1.78	0.89	$\lesssim 18$	$\lesssim 19$
UDS	19627	02:18:17.06	-05:21:38.83	300	21.40	20.91	20.65	20.19	1.37	0.68	$\lesssim 30$	$\lesssim 29$
UDS	29410	02:17:51.22	-05:16:21.84	120	20.59	20.18	19.81	19.36	1.61	0.95	$\lesssim 30$	$\lesssim 22$

common at low redshift, spectroscopic studies become much more difficult at higher redshift as the bulk of the light, and absorption features used to measure kinematics, shift redward into the near-infrared (NIR) (e.g. Kriek et al. 2009; van Dokkum et al. 2009).

New technology such as the new red arm of the LRIS spectrograph at Keck (working beyond $1 \mu\text{m}$) makes it possible to measure velocity dispersions up to $z \sim 1.5$ (Newman et al. 2010, Bezanson et al. 2012). Deep NIR spectroscopy is, however, required to push stellar kinematic studies to even higher redshift. From a ~ 29 hr spectrum of an ultra-compact galaxy at $z = 2.2$ obtained with Gemini Near-IR Spectrograph (Kriek et al. 2009), van Dokkum et al. (2009) found a high, though uncertain, velocity dispersion of $\sigma = 510^{+165}_{-95} \text{ km s}^{-1}$. Onodera et al. (2012) used the MOIRCS on the Subaru telescope to observe the rest-frame optical spectrum of a less-compact, passive, ultra-massive galaxy at $z = 1.82$, but the low spectral resolution and signal-to-noise severely limited the accuracy of their velocity dispersion: $\sigma = 270 \pm 105 \text{ km s}^{-1}$. X-Shooter (D’Odorico et al. 2006), the new UV to NIR spectrograph at the VLT, can provide the required S/N and resolution. The capabilities of X-Shooter for this kind of measurements were demonstrated in van de Sande et al. (2011), who found $294 \pm 51 \text{ km s}^{-1}$ for a massive quiescent galaxy at $z=1.8$. Toft et al. (2012) also use X-Shooter and present a dynamical measurement of a galaxy at redshift $z=2.04$ with similar results. Taken all together, these results indicate that the dynamical and stellar masses are consistent with $z \sim 0$. With the small number of measure-

ments beyond $z > 1.5$, however, the sample is still too small to draw any firm conclusions on whether the stellar masses are truly reliable.

Here we present a sample of 5 massive quiescent galaxies with high Signal-to-Noise (S/N), medium-resolution, UV-NIR spectra at $1.4 < z < 2.1$ observed with X-Shooter on the VLT. The main goal of this paper will be to test if the stellar mass measurements at high redshift are reliable.

The paper is organized as follows. In §2 we present our sample of high-redshift galaxies, the photometric and spectroscopic data, and describe our data reduction. In §3 we determine structural properties and stellar populations, and derive stellar and dynamical masses. We complement our results with stellar kinematic results from other studies at low and high redshift in §4. In §5 we compare our dynamical to the stellar masses. In §6 we study the structural evolution of high-redshift quiescent massive galaxies. In §7 we compare our results with previous measurements and hydrodynamical simulations. Finally, in §8 we summarize our results and conclusions. Throughout the paper we assume a ΛCDM cosmology with $\Omega_m=0.3$, $\Omega_{\Lambda} = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. All broadband data are given in the AB-based photometric system.

2. DATA

2.1. Target Selection

The galaxies in this paper are drawn from the NMBS-I (Whitaker et al. 2010) and UDS (Williams et al. 2009). They were selected to be bright in the H -band, and to have $z > 1.4$, in order to obtain sufficient S/N in 2 hours. The SED from

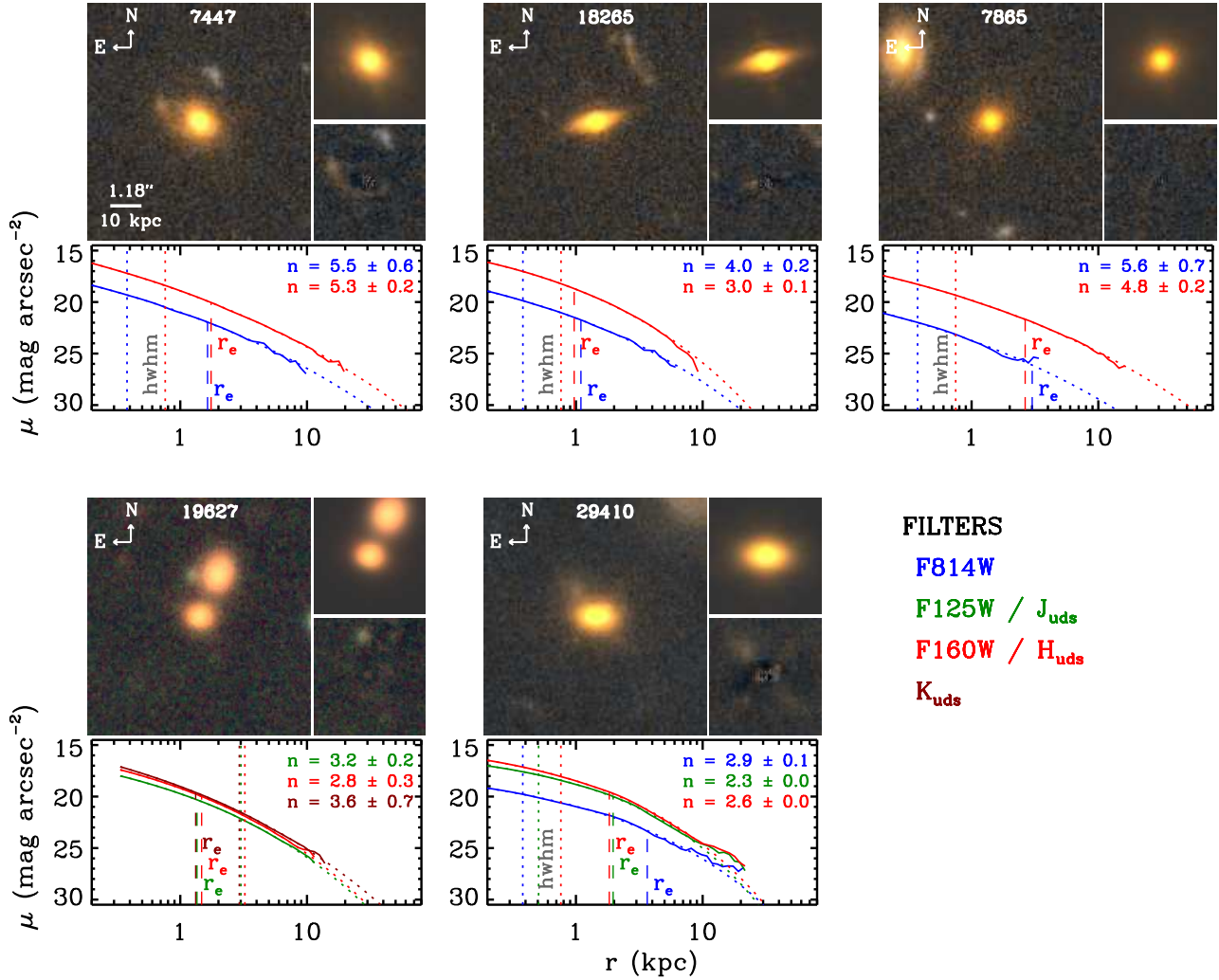


FIG. 2.— Color Images of our five spectroscopic targets. Except for UDS-19627, all galaxies have available HST-WFC3 Imaging. For each target we show the composite color image on the left side, the best Sérsic model from GALFIT and the residual after we subtract the best-fitting model from the original image on the right side. The lower panel shows the intrinsic surface brightness profile with all available bands. Different colors show the different filters, as indicated on the bottom right. Vertical dashed lines show the effective radii for each profile, while the dotted lines shows the FWHM/2 of the PSF. We find color gradients, such that the redder bands have smaller effective radii, for all galaxies but NMBS-COS7447. In that case, the sizes are similar within the errors, but this could be caused by the extra flux of the red arc-like feature in the South-East.

the broad and medium-band photometry was required to indicate that they have quiescent stellar populations, and the rest-frame optical imaging could not show signs of large disturbance due to e.g. mergers. We note that NMBS-COS7447 is presented in van de Sande et al. (2011), and UDS-19627 is presented in Toft et al. (2012). All data for both galaxies have been re-analyzed according to the following procedure for consistency. Our selection had no priors on mass or size, but could be biased in either one of these parameters.

To investigate this, we compare our targets to a sample of galaxies with mass $> 10^{10.5} M_{\odot}$ at $1.4 < z < 2.1$ from the NMBS-I and UDS. Rest-frame $U - V$ and $V - J$ colors are often used to distinguish between star-forming and quiescent galaxies at this redshift (e.g. Williams et al. 2009). Figure 1a shows the UVJ-diagram for all galaxies at redshifts between $1.4 < z < 2.1$ with mass $> 10^{10.5} M_{\odot}$, together with the sample presented here. The size of the squares indicates the density of galaxies. For our targets, the rest-frame colors have been measured from the spectra and best-fitting models, while for the full sample rest-frame colors are based on the broad and medium-band data. As demonstrated by

Williams et al. (2009), non-star-forming galaxies can be selected using a color selection indicated by the black lines. Within this selection region, our targets fall in the region occupied by young, quiescent galaxies (Whitaker et al. 2012). The median specific star formation rates (sSFR), as indicated by the different colors, are in good agreement with the full high-redshift sample at the same place in the UV-VJ diagram. For their mass, however, NMBS-COS7447 and UDS-19627 have slightly bluer colors as compared to the full sample (Figure 1b). At fixed mass, the targets are among the brightest galaxies, except for NMBS-I-7865 (1c). This may not come as a surprise as they are among the youngest quiescent galaxies, and thus have relatively low M/L .

2.2. Spectroscopic Observations

Observations were performed with X-Shooter on the VLT UT2 (D’Odorico et al. 2006). X-Shooter is a second generation instrument on the VLT that consists of 3 arms: UVB, VIS, and NIR. The wavelength coverage ranges from 3000 to 24800 Å in one single exposure. The galaxies were observed in both visitor and service mode, and the observations were

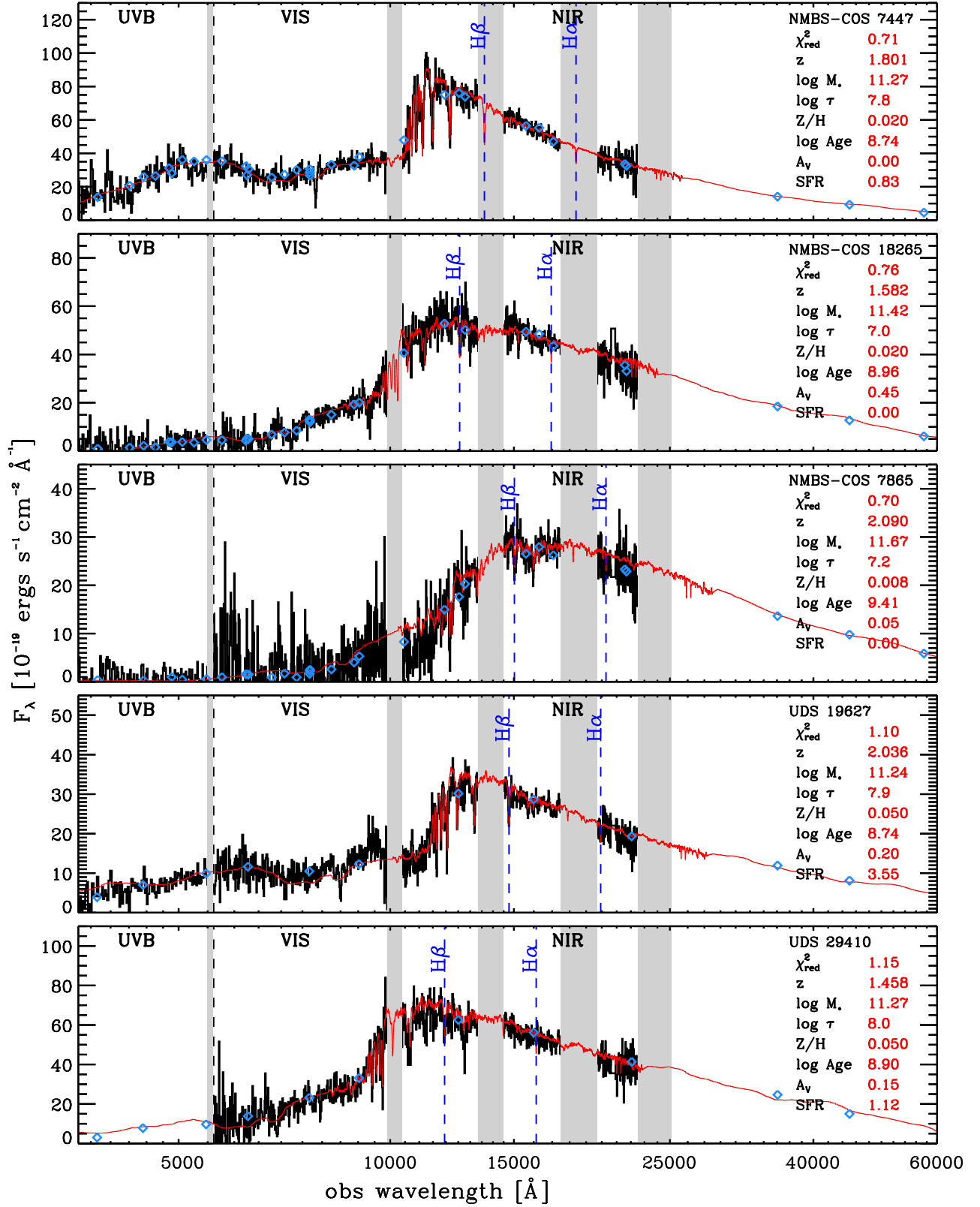


FIG. 3.— UV to NIR X-Shooter spectra in combination with medium- and broad-band data (blue diamonds). The binned spectra ($\sim 10\text{\AA}$) are shown in black, together with the best-fitting BC03 τ -model as shown in red. Grey areas indicate regions with strong sky emission or atmospheric absorption. The UVB spectrum is missing for UDS-29410, due to an instrument problem during the observations. The good agreement between the BC03 models and the spectroscopic data over this large wavelength range is uncanny, and shows the power of modern spectrographs. Both NMBS-COS7865 and UDS-19627 show a small deviation from the best-fitting model around $1\mu\text{m}$, which is caused by the absence of good telluric calibrators. From stellar population synthesis modeling, we find a variety of ages that range from 0.5-2 Gyr. We find no emission lines, and other signs of star formation, and with very little to no dust, these galaxies are quiescent (see Section 3.2).

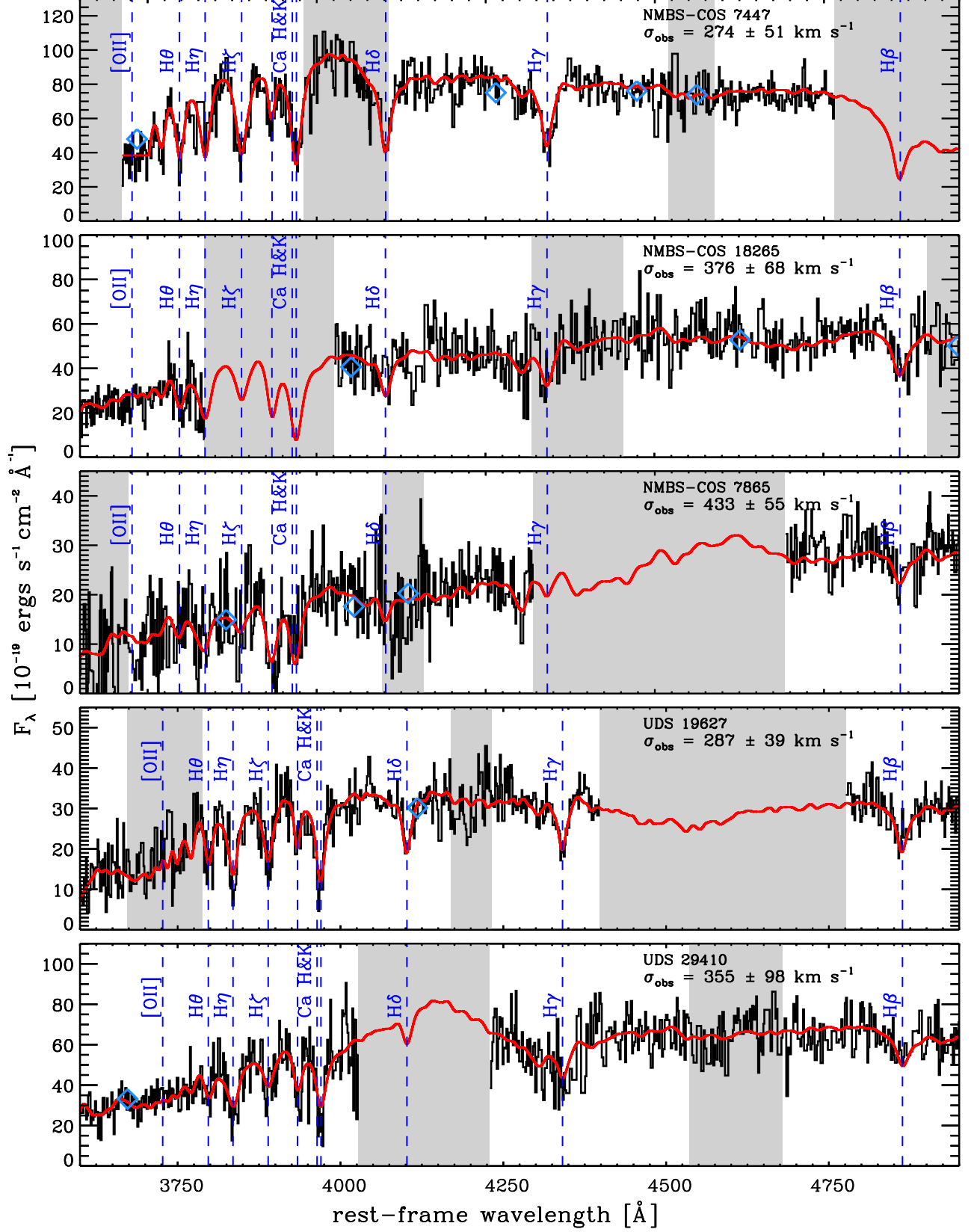


FIG. 4.— Rest-frame optical part of the spectrum focused on the Balmer break. As in Figure 3, the X-Shooter spectrum is shown in black, but this time in higher resolution ($\sim 4\text{\AA}$ observed, or $\sim 100 \text{ km s}^{-1}$ rest-frame). The most prominent absorption features are indicated by the blue dashed lines. The clear detection of absorption lines enables us to measure stellar velocity dispersions. We use pPXF to fit the best-fitting BC03 τ model to the spectrum and find velocity dispersions that range from 275–435 km s^{-1} . The convolved best-fit BC03 template is shown in red.

TABLE 2
STELLAR POPULATION SYNTHESIS PROPERTIES

Catalog	ID	z_{phot}	z_{spec}	$\log \tau$ (yr)	$\log Age$ (yr)	Z	A_V (mag)	$\log M_*$ (M_\odot)	$\log SFR$ ($M_\odot \text{yr}^{-1}$)	$\log sSFR$ (yr^{-1})
NMBS-COS	7447	1.71 ± 0.03	1.800	7.80	8.74	0.020	0.00	11.27	-0.08	-11.35
NMBS-COS	18265	1.60 ± 0.03	1.583	7.00	8.96	0.020	0.45	11.42	-99.00	-99.00
NMBS-COS	7865	2.02 ± 0.05	2.091	7.20	9.41	0.008	0.05	11.67	-99.00	-99.00
UDS	19627	1.94 ± 0.06	2.036	7.90	8.74	0.050	0.20	11.24	0.55	-10.68
UDS	29410	1.44 ± 0.02	1.456	8.00	8.90	0.050	0.15	11.27	0.05	-11.22

NOTE. — No errors are provided, as the 68% confidence all fall within one grid point. The real errors are dominated by systematic uncertainties.

carried out between January 2010 and March 2011 (Programs: Fynbo 084.A-0303(D), Van de Sande 084.A-1082(A), Franx 085.A-0962(A), Toft 086.B-0955(A)). Full information on the targets and observations are listed in Table 1. All observations had clear sky conditions and an average seeing of 0.8". A 0.9" slit was used in the NIR, except for the 1st hour of UDS-19627 where a 0.6" slit was used. For the 0.9" slit, this resulted in a spectral resolution of 5100 at $1.4\mu\text{m}$. Observing blocks were split in exposures of each 10 to 15 minutes with an $ABA'B'$ on-source dither pattern. For most targets, a telluric standard of type B8V-B9V was observed before and after our primary target, in order to create a telluric absorption spectrum at the same airmass as the observation of our target.

2.3. Spectroscopic Reduction

Data from the three arms of X-Shooter must be analyzed separately and then combined to cover the full range from the UV-NIR. In the NIR we identified bad pixels in the following way. The data were corrected for dark current, flatfielded, and sky subtracted using the average of the preceding and subsequent frames. The ESO pipeline (version 1.3.7, Goldoni et al. 2006) was used to derive a wavelength solution for all orders. The orders were then straightened using integer pixel shifts to retain the pixels affected by cosmic-rays and bad pixels. Additional sky subtraction was done on the rectified orders, by subtracting the median in the spatial direction. Cosmic rays and bad pixels were identified by LA-Cosmic (van Dokkum 2001), and a bad pixel mask was created.

Further 3-sigma clipping was done on the different exposures, corrected for dithers, to identify any remaining outliers. The bad pixel masks of different orders were combined into a single file and then transformed back to the raw frame for each exposure. In the following, the mask file will follow the same rectification and wavelength calibration steps as the science frames.

The raw science frames were corrected for dark current, flatfielded, and sky subtracted using the average of the preceding and subsequent frames. Next the observations were rectified and wavelength calibrated, only this time we used interpolation when rectifying the different orders. Again, additional sky subtraction was done. Per order, all exposures were combined, with exclusion of bad pixels and those contaminated with cosmic rays present in the mask file.

The telluric spectra were reduced in the same way as the science frames. We constructed a response spectrum from the telluric stars in combination with a stellar model for a B8V/B9V star from a black body curve and models from Munari et al. (2005). Residuals from Balmer absorption features in the spectrum of the tellurics were removed by interpolation. All the orders of the science observations were cor-

rected for instrumental response and atmospheric absorption by dividing by the response spectrum.

The different orders were then combined, and in regions of overlap weighted using the S/N of the galaxy spectrum. A noise spectrum was created by measuring the noise in the spatial direction below and above the galaxy. If the regions exceeded an acceptable noise limit, from contamination by OH lines or due to low atmospheric transmission, this spatial region was discarded for further use. The 2D spectra were visually inspected for emission lines, but none were found. A 1D spectrum was extracted by adding all lines (along the wavelength direction), with flux greater than 0.1 times the flux in the central row, using optimal weighting. Absolute flux calibration was performed by scaling the spectrum to the available photometric data.

A low resolution spectrum was constructed by binning the 2D spectrum in wavelength direction. Using a bi-weight mean, 20 good pixels, i.e. not affected by skylines or strong atmospheric absorption, were combined. The 1D spectrum was extracted from this binned 2D spectrum in a similar fashion as the high resolution spectrum (*see Figure 3 & 4*).

For the UVB and VIS arms we used the ESO pipeline (version 1.3.7, Goldoni et al. 2006) to correct for the bias, flat-field, and dark current, and to derive the wavelength solution. The science frames were also rectified using the pipeline, but thereafter, treated in exactly the same as the rectified 2D spectra of the NIR arm, as described above.

2.4. Imaging

Four different imaging datasets are used to measure the surface brightness profiles of our galaxies, as summarized below. (1) All our target in the NMBS-I field were observed with HST-WFC3 H_{160} as part of the program HST-GO-12167 (PI: Franx). Each targets was observed for one orbit (2611 sec), using a four point dither pattern, with half pixel offsets. Reduction of the data was done in a similar way as to that described in Bouwens et al. (2010), but without sigma-clipping in order to avoid masking the centers of stars. The drizzled images have a pixel scale of 0.06", with a full width at half-maximum (FWHM) of the point-spread function (PSF) of $\sim 0.16''$. (2) Our NMBS-I targets are complemented with HST-ACS I_{814} imaging from COSMOS (v2.0, Koekemoer et al. 2007; Massey et al. 2010), which has a 0.03" pixel scale and PSF-FWHM of $\sim 0.11''$. (3) For UDS-29410 we make use of the HST-ACS $F814W$, HST-WFC3 J_{125} and H_{160} from UDS-CANDELS (Grogin et al. 2011; Koekemoer et al. 2011). These data have the same properties as the data described in (1) and (2). (4) For UDS-19627 we use ground based data from UKIDSS-UDS, (Lawrence et al. 2007; Warren et al. 2007) Data Release 8 in the J, H, and K-

TABLE 3
COMPILATION OF MASSES AND STRUCTURAL PARAMETERS FOR HIGH-REDSHIFT GALAXIES

Reference ^a	ID	z_{spec}	r_e	n_{seric}	b/a	σ_e	σ_e/σ_{ap}	$\beta(n)$	$\log M_{\text{dyn}}$	$\log M_{*,\text{corr}}^c$	Filter
0	7447	1.800	1.75 ± 0.21	5.27 ± 0.23	0.71 ± 0.02	287^{+55}_{-52}	1.048	5.16	$11.24^{+0.13}_{-0.13}$	11.22	H_{F160W}
0	18265	1.583	0.97 ± 0.12	2.97 ± 0.06	0.26 ± 0.02	400^{+78}_{-66}	1.065	6.61	$11.38^{+0.13}_{-0.11}$	11.32	H_{F160W}
0	7865	2.091	2.65 ± 0.33	4.82 ± 0.15	0.83 ± 0.02	446^{+54}_{-59}	1.031	5.42	$11.82^{+0.09}_{-0.09}$	11.64	H_{F160W}
0	19627	2.036	1.32 ± 0.17	3.61 ± 0.73	0.48 ± 0.06	304^{+43}_{-39}	1.059	6.18	$11.24^{+0.10}_{-0.09}$	11.20	K
0	29410	1.456	1.83 ± 0.23	2.59 ± 0.03	0.54 ± 0.02	371^{+114}_{-90}	1.045	6.88	$11.61^{+0.19}_{-0.15}$	11.24	H_{F160W}
...

NOTE. — This Table will be published in its entirety in the electronic edition of ApJ, and can also be downloaded from <http://www.strw.leidenuniv.nl/~sande/data/>. A portion is shown here for guidance regarding its form and content.

^a References: 0) This Work 1) Bezanson et al. (2012); 2) van Dokkum et al. (2009); 3) Onodera et al. (2012); 4) Cappellari et al. (2009); 5) Newman et al. (2010); 6) van der Wel et al. (2008) & Blakeslee et al. (2006) 7) Toft et al. (2012).

^b Stellar masses are corrected to account for the difference between the catalog magnitude and our measured magnitude.

band, as no as HST data is available. Imaging from all three bands were drizzled to a pixel scale $0.134''$, and the FWHM of the PSF is $0.7''$ in the K-band.

3. STRUCTURAL PROPERTIES AND STELLAR POPULATIONS

3.1. Surface Brightness Profiles

Radial profiles are measured for all galaxies on all available imaging as described in Section 2.4. Galaxies are fitted by two-dimensional Sérsic radial surface brightness profiles (Sérsic 1968), using GALFIT (version 3.0.2; Peng et al. 2010). Relatively large cutouts of $25'' \times 25''$ were provided to GALFIT to ensure an accurate measurement of the background, which was a free parameter in the fit. All neighboring sources were masked using a segmentation map obtained with SExtractor (Bertin & Arnouts 1996). In the case of UDS-19627, the close neighbor was fitted simultaneously. Bright unsaturated field stars were used for the PSF convolution. All parameters, including the sky, were left free for GALFIT to determine.

Even though galaxies at low redshift are well-fitted by single Sérsic profiles (e.g. Kormendy et al. 2009), this does not necessarily have to be true for galaxies at $z \sim 2$. Therefore, we correct for missing flux using the method described in Szomoru et al. (2010). We find very small deviation in residual corrected effective radii, with a median absolute deviation of 3.4%. Color images and measured profiles are shown in Figure 2.

We repeated the measurements using a variety of PSF stars ($N \sim 25$). We find an absolute median deviation in the half-light radius of $\sim 3\%$ for HST-WFC3, $\sim 3.5\%$ for HST-ACS, and $\sim 10\%$ for the ground-based UDS-UKIDSS data, due to variations in the PSF. The largest source of uncertainty in the measured profiles is, however, caused by the error in the sky background estimate. Even though these galaxies are among the brightest at this redshift, using the wrong sky value can result in large errors for both r_e and n . We determine the error in the sky background estimate by measuring the variations of the residual flux in the profile between 5 and 15 arcseconds. For sizes derived from HST-ACS, the absolute median deviation in the effective radius due to the uncertainty in background is $\sim 13\%$, and for HST-WFC3 $\sim 12\%$. Due to the deeper ground-based UDS-UKIDSS data, the uncertainty for UDS-19627 due to the sky is $\sim 8\%$. All of our results are summarized in Table 3. In what follows, we will use the mean effective radius and Sérsic n from the band which is closest to rest-frame optical r' .

3.2. Stellar population properties

We estimate the stellar population properties by fitting the low-resolution ($\sim 10\text{\AA}$ in observed frame) spectrum in the Visual and NIR in combination with the broad- and medium-band photometry with SPS models. We exclude the UVB part of the spectrum due to the lower S/N and the extensive high S/N broad band photometry in this wavelength region. Stellar templates from Bruzual & Charlot 2003 (BC03) are used, with an exponentially declining star formation history with timescale τ , together with a Chabrier (2003) IMF, and the Calzetti et al. (2000) reddening law.

Using the FAST code (Kriek et al. 2009) we fit a full grid in age, dust content, star formation timescale, and metallicity. We adopt a grid for τ between 10 Myr and 1 Gyr in steps of 0.1 dex. The age range can vary between 0.1 Gyr and 10 Gyr, but the maximum age is constrained to the age of the Universe at that particular redshift. Step size in age is set as high as the BC03 templates allow, typically 0.01 dex. Metallicity can vary between $Z = 0.004$ (subsolar), $Z = 0.08$, $Z = 0.02$ (solar), and $Z = 0.05$ (supersolar). Furthermore, we allow dust attenuation to range from 0 to 2 magnitudes with step size of 0.05. The redshift used here is from the best-fitting velocity dispersion (see Section 3.3). Results are summarized in Table 2.

Due to our discrete grid and the high quality data, and also because metallicity and age are limited by the BC03 models, our 68% confidence levels are all within one grid point. Our formal errors are therefore mostly zero, and not shown in Table 2. This does not reflect the true uncertainties, which are dominated by choice of SPS models, IMF, SFH, and extinction law (see e.g. Conroy et al. 2009; Muzzin et al. 2009).

The low sSFR confirms the quiescent nature of the galaxies in our sample, and they match well with the sSFR of the general population in the same region of the UVJ diagram (Figure 1a). We find a range of metallicities, with the oldest galaxy having the lowest metallicity. However, due to the strong dependency between age and metallicity, we do not believe this result to be significant. Overall, the dust content in our galaxies is low.

The galaxies in our sample are not detected at $24\text{ }\mu\text{m}$, leading to a $3\text{-}\sigma$ upper-limit of $18\text{ }\mu\text{Jy}$ for the galaxies in NMBS-COSMOS, and $30\text{ }\mu\text{Jy}$ in the UDS (see Whitaker et al. 2012; Toft et al. 2012). From these upper limits we calculate the dust-enshrouded SFRs which are listed in Table 1.

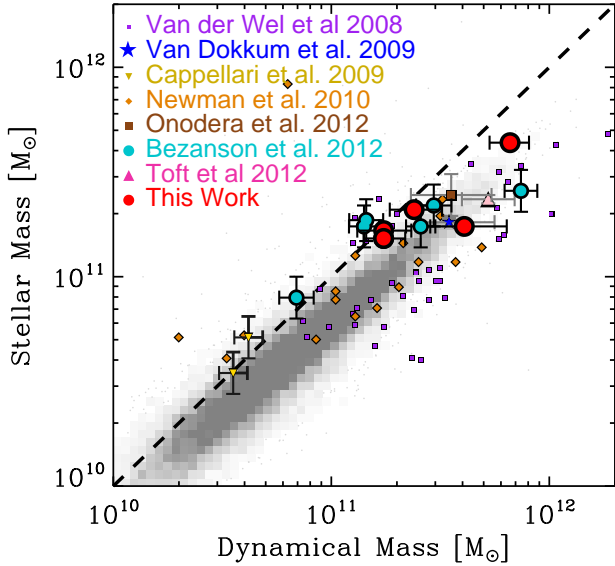


FIG. 5.— Comparison of the stellar mass versus the dynamical mass. Grey squares are again non-star-forming galaxies from the SDSS. Different symbols are from a compilation of high-redshift galaxies as described in Section 4.2. The dashed line is for equal dynamical and stellar mass. Low-redshift galaxies are all below the line, as is expected from the contribution of dark matter. All our high-redshift galaxies have dynamical masses that are equal or larger than the stellar mass. This suggests that the stellar mass measurements at high-redshift are robust for passive galaxies.

3.3. Velocity dispersions

The clear detection of absorption lines in our spectra, together with the medium resolution of X-Shooter, allow us to measure accurate stellar velocity dispersions. We use the unbinned spectra in combination with the Penalized Pixel-Fitting method (pPXF) by Cappellari & Emsellem (2004) and our best-fitting BC03 models as templates. Spectra were put on to a logarithmic wavelength scale without using interpolation, but with masking of the bad pixels. The effect of template mismatch was reduced by simultaneously fitting the template with a ~ 17 -order Legendre Polynomial. Our results depend only slightly on the choice of the order of the polynomial (Appendix A). Together with the measured velocity dispersion, the fit also gives us the line of sight velocity, and thus z_{spec} .

In particular for the younger galaxies in our sample that show a clear signature of A-type stars, we find a dependence of the measured velocity dispersion as a function of template age. A more stable fit is obtained when restricting the wavelength range to $4020\text{\AA} < \lambda < 7000\text{\AA}$, which excludes the Balmer break region (see also Appendix A).

The errors on the velocity dispersion were determined in the following way. We subtracted the best-fit model from the spectrum. The residuals were randomly rearranged in wavelength space and added to the best-fit template. We determined the velocity dispersion of 500 simulated spectra. Our quoted error is the standard deviation of the resulting distribution of the measured velocity dispersions. When we include the Balmer break region in the fit, the formal random error decreases, but the derived dispersion becomes more dependent on the chosen stellar template. In total we have three high-quality measurements, and two with medium quality. We note that if we exclude the two medium-quality measurements from our sample, our science results would not change.

All dispersions are corrected for instrumental resolution ($\sigma=23 \text{ km s}^{-1}$) and the spectral resolution of the templates

($\sigma=85 \text{ km s}^{-1}$). Furthermore, we apply an aperture correction to our measurements as if they were observed within a circular aperture radius of r_e . In addition to the traditional correction for the radial dependence of velocity dispersion (e.g. Cappellari et al. 2006), we account for the effects of the non-circular aperture, seeing and optimal extraction of the 1-D spectrum. The aperture corrections are small with a median of 4.8% (See Appendix B). The final dispersions and corresponding uncertainties are given in Table 3.

3.4. Dynamical Masses

Combining the size and velocity dispersion measurements we are now able to estimate dynamical masses using the following expression:

$$M_{\text{dyn}} = \frac{\beta(n) \sigma_e^2 r_e}{G}. \quad (1)$$

Here $\beta(n)$ is an analytic expression as a function of the Sérsic index, as described by Cappellari et al. (2006):

$$\beta(n) = 8.87 - 0.831n + 0.0241n^2. \quad (2)$$

This is computed from theoretical predictions for β from spherical isotropic models described by the Sérsic profile, for different values of n , and integrated to one r_e (cf. Bertin et al. 2002). Using a Sérsic depended virial constant $\beta(n)$ gives a better correspondence between M_{dyn} and M_* for galaxies in the SDSS (Taylor et al. 2010b). This does require however, that the total stellar masses are also derived using the luminosity of the derived Sérsic profile. Thus we correct our total stellar mass, as derived from the total magnitude as given in the catalogs (measured with SExtractor), to the total magnitude from the Sérsic fit. We note that the values for β that we find are all close to 5, a value often used in the literature (e.g. Cappellari et al. 2006). Our dynamical masses and corrected stellar masses are given in Table 3.

4. COMPILATION OF KINEMATIC STUDIES

Before we are able to study the structural evolution of our galaxies, we first combine the results from different kinematic studies at various redshifts. Where possible, we apply similar corrections as described above.

4.1. Low-redshift sample

At low redshift we select galaxies from the Sloan Digital Sky Survey (SDSS, DR7). Stellar masses are based on MPA⁸ fits to the photometry following the philosophy of Kauffmann et al. (2003), and Salim et al. (2007). Star formation rates are based on Brinchmann et al. (2004). Structural parameters are from the NYU Value-Added Galaxy Catalog (NYU-VAGC, Blanton et al. 2005). For all galaxies, velocity dispersions were aperture corrected as described in Section 3.3, and stellar masses are calculated with a Chabrier (2003) IMF. We furthermore correct the stellar masses using the total magnitude from the best Sérsic fit. All dynamical masses were derived using Equation 1. For making an accurate comparison between low- and high-redshift galaxies, we only select non-star-forming galaxies, i.e. $s\text{SFR} < 0.3/t_H$ (see Williams et al. 2009), where t_H is the age of the universe at the given redshift.

⁸ <http://www.mpa-garching.mpg.de/SDSS/DR7/>

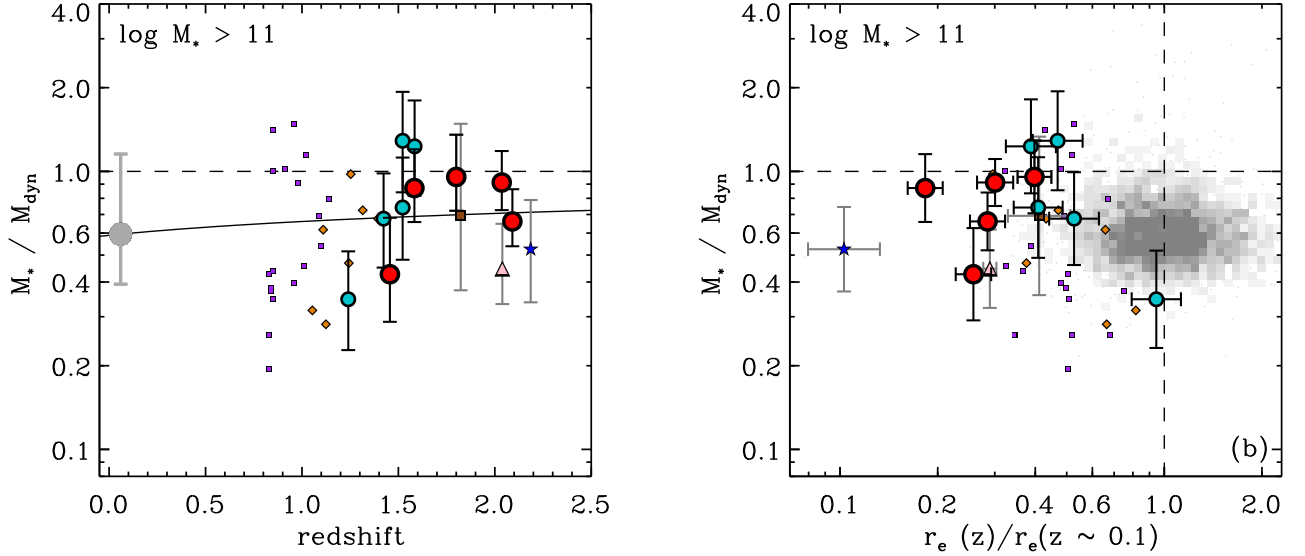


FIG. 6.— *Left panel:* Stellar mass divided by the dynamical mass versus redshift. Galaxies below the line have dynamical mass greater than the stellar mass, above the line is the non-physical regime. For SDSS galaxies with stellar mass $> 10^{11} M_\odot$, we find that they have a median M_*/M_{dyn} of 0.59. Up to redshift $z = 1.5$ we find a similar slightly lower median value (~ 0.5), but it rapidly increases at $z > 1.5$ with median a of $M_*/M_{\text{dyn}} = 0.9$. The solid line is the best-fit $M_*/M_{\text{dyn}} \propto (1+z)^{0.17 \pm 0.04}$. We caution that this result might be biased due to the selection effects as explained in Section 3, and relatively large measurement errors. *Right panel:* Stellar mass divided by the dynamical mass versus the evolution in the effective radius at fixed dynamical mass. Galaxies that have small effective radii at fixed dynamical mass also show higher ratios of M_*/M_{dyn} , although there is a significant scatter.

4.2. Intermediate- and high-redshift sample

Our high redshift sample consists of a collection of both optical and NIR spectroscopic studies on non-stacked samples. van der Wel et al. (2008) present a sample of quiescent galaxies at $z \sim 1$, which itself is a compilation of three studies in the following fields: Chandra Deep Field South (CDF-S; van der Wel et al. 2004; 2005), the Hubble Deep Field North (HDF-N; Treu et al. 2005a; 2005b), and cluster galaxies in MS 1054-0321 at $z = 0.831$ (Wuyts et al. 2004). We derive stellar masses for this sample by running FAST on available catalogs, i.e., FIREWORKS (Wuyts et al. 2008) for the CDS-S, Skelton et al. (in prep.) for the HDF-N, and FIRES (Förster Schreiber et al. 2006) for MS 1054-0321. For CDF-S and HDF-N the stellar masses are corrected using the total magnitude from the best $n = 4$ fit to be consistent with the structural parameters from van der Wel et al. (2008). For MS 1054-0321, we use structural parameters and stellar mass corrections based on the results by Blakeslee et al. (2006), who fit Sérsic profiles with n as a free parameter.

Other high-redshift results included here are from Newman et al. (2010), and Bezanson et al. (2012), who use the new red-arm of LRIS on Keck to obtain UV rest-frame spectra of galaxies at $z \sim 1.3$ and $z \sim 1.5$ respectively. Velocity dispersions for two galaxies at $z = 1.41$ are presented by Cappellari et al. (2009), and have been observed with VLT-FORS2 (but see also Cenarro & Trujillo 2009). Using NIR spectrographs, Onodera et al. (2012, Subaru-MOIRCS), and van Dokkum et al. (2009, GNIRS) obtained velocity dispersions for two galaxies at $z = 1.82$ and $z = 2.186$. Similar to this study, Toft et al. (2012) study UDS-19627 using VLT X-Shooter. Dynamical masses were derived using Equation 1. Note that for the studies of Cappellari et al. (2009), Onodera et al. (2012), van Dokkum et al., and Toft et al. (2012) no stellar mass corrections were applied due to the absence of the necessary information. All structural and kinematic properties of our high-redshift sample are listed in Table 3.

5. ARE STELLAR MASSES RELIABLE?

The main goal of this paper is to see whether the stellar masses at $z \sim 2$ are reliable. Here we compare our stellar masses, as derived from the spectra and photometry, to our dynamical masses, which are derived from effective radii and stellar velocity dispersions (Figure 5). Grey squares represent the density of non-star-forming, low-redshift galaxies from the SDSS as described in Section 4.1. Other symbols are the high-redshift studies as described in Section 4.2. The one-to-one relation for M_{dyn} and M_* is indicated by the dashed line. Note that the region above the line is nonphysical with stellar masses being higher than the dynamical mass.

We find that most $z > 1.5$ galaxies are very massive, in the range $11.2 < \log M_{\text{dyn}}/M_\odot < 11.8$. At all redshifts, stellar and dynamical masses are tightly correlated and dynamical mass, which includes baryonic and dark matter, is always higher than stellar mass. Thus, we infer that the stellar masses are broadly correct, and that the apparent size evolution of massive galaxies in photometric studies cannot be explained by errors in the photometric masses (see also van der Wel et al. 2008).

Figure 6a shows the ratio of the stellar and dynamical mass as function of redshift for all galaxies with stellar mass $> 10^{11} M_\odot$. We see that the average ratio at low-redshift for massive galaxies is a factor of 0.59 with a scatter of 0.12 dex. We note that For MS 1054-0321, the ratio of the stellar to dynamical mass are slightly higher as compared to low redshift galaxies. Up to redshift $z \sim 1.5$ we find a similar value (~ 0.5) with similar scatter, but at higher redshift, the ratio seems to decline. For galaxies at $z > 1.5$ we find a median ratio of $M_*/M_{\text{dyn}} = 0.9$. We quantify the evolution in this ratio by fitting the relation:

$$M_*/M_{\text{dyn}} \propto (1+z)^\alpha. \quad (3)$$

We find $\alpha = 0.17 \pm 0.04$ (linear fit in log-log space), shown as the solid black line in Figure 6a. Due to the relatively large measurements errors as compared to low redshift, and the selection bias, we are cautious to draw any strong conclusions

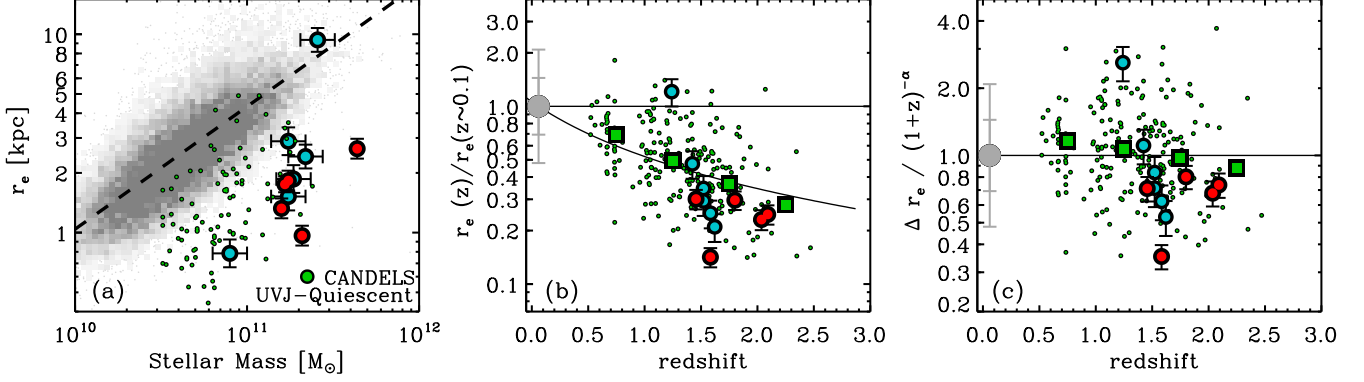


Fig. 7.— Structural comparison of our spectroscopic sample to the full population. *Left panel:* Effective radii versus mass for low- and high-redshift galaxies. Grey squares are non-star-forming $z \sim 0$ galaxies from the SDSS, with the dashed line the best-fit to Equation 4. High-redshift quiescent galaxies ($1.4 < z < 2.1$) from CANDELS are shown as green circles, together with our spectroscopic sample shown as red circles, and the sample by Bezanson et al. (2012) shown as cyan circles. At fixed mass, our high-redshift galaxies have smaller effective radii, similar to what has been found by many studies. *Middle panel:* Evolution in effective radius at fixed stellar mass, thus corrected for the $M_* - r_e$ relation from the left panel. Big green squares are the median effective radii in bins of redshift for the CANDELS data. The solid line is the best $r_e \propto (1+z)^{-1.02 \pm 0.05}$. At similar redshift, we find that our sample and that of Bezanson et al. (2012) is mostly below this fit, indicating that our samples are biased towards smaller effective radii. *Right panel:* Similar to the middle panel, but now divided by $(1+z)^{-1.02}$ for a better comparison of our spectroscopic sample to the CANDELS data. When comparing the median of our sample to the binned median of other quiescent galaxies at similar redshift, we find smaller effective radii by a factor ~ 1.17 . This might be explained by us using a selection based on aperture magnitude, which tends to be biased towards smaller galaxies.

from this. Another way of thinking about this is that the velocity dispersions inferred from the stellar mass and effective radii are higher than the measured velocity dispersions (see also Bezanson et al. 2012). It is tempting to speculate that this might be caused by a decrease in the dark matter fraction as a function of redshift. If galaxies grow in size over time, then the dark matter fraction within r_e will also increase in a similar fashion as shown here (e.g. Hopkins et al. 2009a). If so, this could also indicate that the IMF at high-redshift is very similar to the IMF at low-redshift.

Figure 6b shows M_*/M_{dyn} versus the evolution of the effective radius at fixed dynamical mass (see Section 6.2 and Figure 8). Although there is significant scatter, we do find the galaxies with high M_*/M_{dyn} also tend to have smaller size at fixed dynamical mass. Galaxies that are closest to the present-day mass-size relation (dashed vertical line) show lower ratios of stellar to dynamical mass. This agrees with the idea that the decreasing ratio of M_*/M_{dyn} with time might be correlated to the size growth of massive galaxies.

6. STRUCTURAL EVOLUTION OF QUIESCENT GALAXIES

From the previous section we now know that the dynamical masses are consistent with the stellar masses. However, for the high-redshift galaxies in our sample, the dynamical masses are somewhat lower than expected from the stellar masses. For this reason, we will re-examine the structural evolution of massive quiescent galaxies but now using dynamical measurements.

6.1. Bias towards compact galaxies

As noted in Section 2.1, this sample is biased towards young quiescent galaxies. Therefore, we will first investigate whether our sample and that of Bezanson et al. (2012) is also biased in size as compared to other high-redshift galaxies. We gathered structural properties of galaxies from two studies that use CANDELS data in the UDS and GOODS-South fields (Patel et al. 2012; Szomoru et al. 2012). We compare to a subsample of these galaxies which are determined to be quiescent from their rest-frame U-V and V-J colors (see e.g. Figure 1a). When comparing the effective radii versus the stellar mass in Figure 7a, we find that our galaxies (red circles) and

those of Bezanson et al. (2012) (cyan circles) are in general more compact as compared to the high-redshift CANDELS galaxies (small green circles).

For low-redshift galaxies we parametrize the mass-size relation by:

$$r_e = r_c \left(\frac{M_*}{10^{11} M_\odot} \right)^b \quad (4)$$

(Shen et al. 2003; van der Wel et al. 2008). Using a linear least-squares fit in log-log space, we find best fitting values of $r_c = 4.32$ kpc and $b = 0.62$. This is slightly different from the Shen et al. (2003) results who find $r_c = 4.16$ kpc and $b = 0.56$. The difference may be explained by different selection criteria, and their use of an older release version of SDSS. Figure 7b shows the evolution in effective radius, by comparing galaxies with similar mass at different redshifts. Using both the SDSS and the CANDELS data, we examine the amount of evolution in size by fitting the following relation:

$$r_e \propto (1+z)^\alpha, \quad (5)$$

We find $\alpha = -1.02 \pm 0.05$ (linear fit in log-log space). Our spectroscopic targets and those of Bezanson et al. (2012) are mostly below this best-fit relation, being smaller by a factor of ~ 1.17 as compared to median in redshift bins (big green squares). This is especially clear from Figure 7c, where we correct for the evolution in size.

The explanation for this bias might be related to the way our targets are selected. As our selection is based on the magnitude within a fixed aperture of $1.5''$, instead of the total magnitude, we create a bias towards compact galaxies. For galaxies with similar total magnitudes, the smaller galaxies will be brighter within a photometric aperture, which will increase the S/N of their spectra, and thus make it into our sample. In what follows, we will correct for this bias by increasing our sizes and those of Bezanson et al. (2012) by a factor of $\sqrt{1.17}$, and decreasing the velocity dispersion by a factor of $\sqrt{1.17}$.

6.2. Evolution in Size

In Figure 8a we plot effective radius versus dynamical mass. Symbols are the same as in Figure 6. For the low-redshift

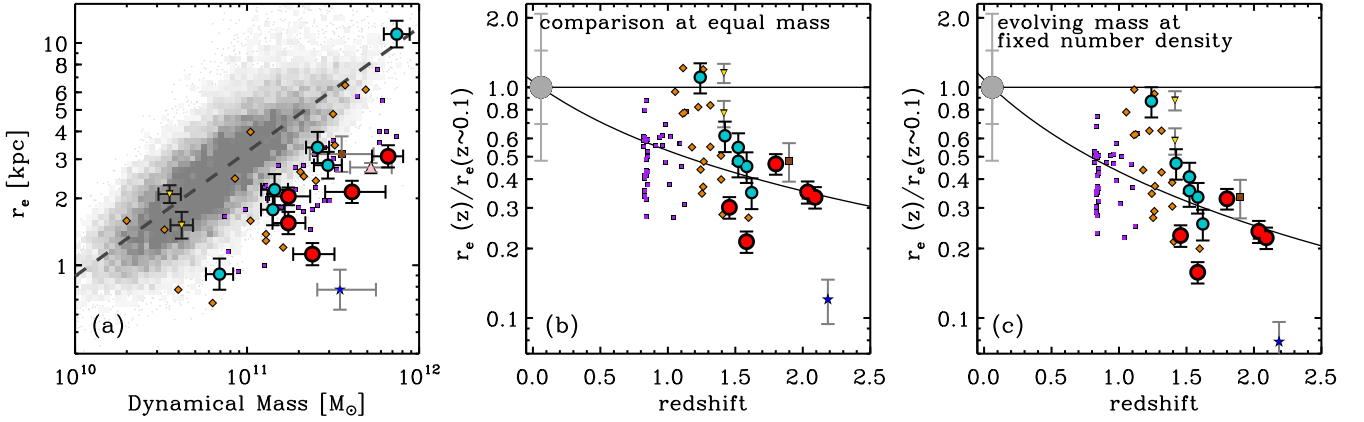


FIG. 8.— Redshift evolution of the effective radius in passive galaxies. *Left panel:* Size versus dynamical mass. Dashed line is the best-fit using Equation 6 for non-star-forming SDSS galaxies. At fixed dynamical mass, we find that our galaxies are smaller by a factor ~ 3 compared to low-redshift galaxies. *Middle panel:* Evolution of the effective radius at fixed dynamical mass versus redshift. The solid line is the best-fit $r_e \propto (1+z)^{-0.99 \pm 0.1}$, similar to what has been found by other kinematic studies at high-redshift. Similar to Figure 7, our sample is below the best-fit line to the entire high-redshift sample. *Right panel:* Evolution of the effective radius using an evolving mass function at constant number density. We now compare galaxies at high-redshift, to more massive galaxies at low-redshift, assuming the mass evolves as $\Delta \log M/M_{\odot} \sim 0.15z$. This time, we find an even stronger evolution, with $r_e \propto (1+z)^{-1.32 \pm 0.1}$.

galaxies we parametrize the mass-size relation according to the following equation

$$r_e = r_c \left(\frac{M_{\text{dyn}}}{10^{11} M_{\odot}} \right)^b, \quad (6)$$

and find $r_c = 3.23$ kpc and $b = 0.56$ (dashed line). This is in good agreement with $b = 0.56$ and $r_c = 3.26$ kpc as found by van der Wel et al. (2008). At fixed dynamical mass, we see that all our galaxies have smaller effective radii as compared to low-redshift. This is further illustrated in Figure 8b, where we compare the effective radii at fixed dynamical mass to the mass-size relation at $z \sim 0$. The solid line is the best-fit as described by equation 5, with $\alpha = -0.99 \pm 0.1$. This is in agreement to what has been found in previous kinematical studies (van der Wel et al. 2008; Newman et al. 2010). The scatter between different studies is considerable, with the work by van Dokkum et al. (2009) having the largest size difference while Onodera et al. (2012) has the smallest. This sample falls in between the two extremes, but we note that corrected for the bias in size (Section 6.1).

Instead of comparing galaxies sizes at fixed dynamical mass, we will now take into account that galaxies do evolve in mass (e.g. Patel et al. 2012). In van Dokkum et al. (2010) they find that, for a sample selected at a constant number density, the stellar mass evolves as

$$\log M_n/M_{\odot} = 11.45 - 0.15z. \quad (7)$$

The number density that this result is based upon, $n = 2 \times 10^{-4} \text{Mpc}^{-3}$, corresponds to an average mass of $\log M_*/M_{\odot} \sim 11.15$ at $z \sim 2$, similar to our sample. Assuming that the mass evolves as $\Delta \log M/M_{\odot} \sim 0.15z$, we will now use mass evolution to compare effective radii for galaxies at different redshifts. For example, a galaxy with $\log M_{\text{dyn}}/M_{\odot} = 11$ at $z \sim 2$ will thus be compared with a $z \sim 0$ galaxy with $\log M_{\text{dyn}}/M_{\odot} = 11.3$. Figure 8c shows the evolution in size at fixed number density as function of redshift. Not surprising, the evolution in effective radii is more extreme, as we are now comparing $z \sim 2$ galaxies to more massive, and therefore bigger galaxies at $z \sim 0$. Again using Equation 5, we now find that $\alpha = -1.32 \pm 0.1$ provides the best fit. In conclusion, assuming that galaxies evolve in both mass and size, we find that the effective radii have to grow by a factor ~ 4 from $z \sim 2$

to the present day.

6.3. Evolution in Velocity Dispersion

In Figure 9a we compare the stellar velocity dispersion within one r_e versus the dynamical mass for both low- and high-redshift samples. The dashed line is the parametrization of the $\sigma_e - M_{\text{dyn}}$ relation for low- z galaxies using the following equation:

$$\sigma_e = \sigma_c \left(\frac{M_{\text{dyn}}}{10^{11} M_{\odot}} \right)^b. \quad (8)$$

We find that $\sigma_c = 148.9 \text{ km s}^{-1}$ and $b = 0.24$. Our high-redshift sample is clearly offset from low-redshift galaxies in the SDSS, i.e. at fixed mass they have higher velocity dispersions. Comparison of the velocity dispersion at fixed dynamical mass, as seen in 9b, shows a clear evolution in σ_e , such that velocity dispersion decreases over time. From this figure the increase in accuracy for the velocity dispersion measurements with new instruments, as compared to other older studies, is also clearly noticeable. Again we use the following simple relation to quantify the amount of evolution:

$$\sigma_e \propto (1+z)^{\alpha}, \quad (9)$$

and find that $\alpha = 0.50 \pm 0.08$. From $z \sim 2$ to $z \sim 0$ the stellar velocity dispersions decrease by a factor ~ 1.7 . Again, we note that we applied a correction to the velocity dispersions in our sample, in order to correct for the bias towards more compact galaxies (Section 6.1). If we now compare low- and high-redshift galaxies using an evolving mass function as described above, we find that the velocity dispersion decreases less than when comparing at fixed dynamical mass. If we use equation 9, we find that $\alpha = 0.36 \pm 0.08$. In other words, the velocity dispersion decreases by a factor ~ 1.4 from $z \sim 2$ to present-day.

6.4. Evolution in Mass Density

Next, we will focus on the central and effective mass densities using a similar approach as described in Saracco et al. (2012). In short, using the intrinsic Sérsic profile we can calculate the fraction of the luminosity that is within 1 kpc as

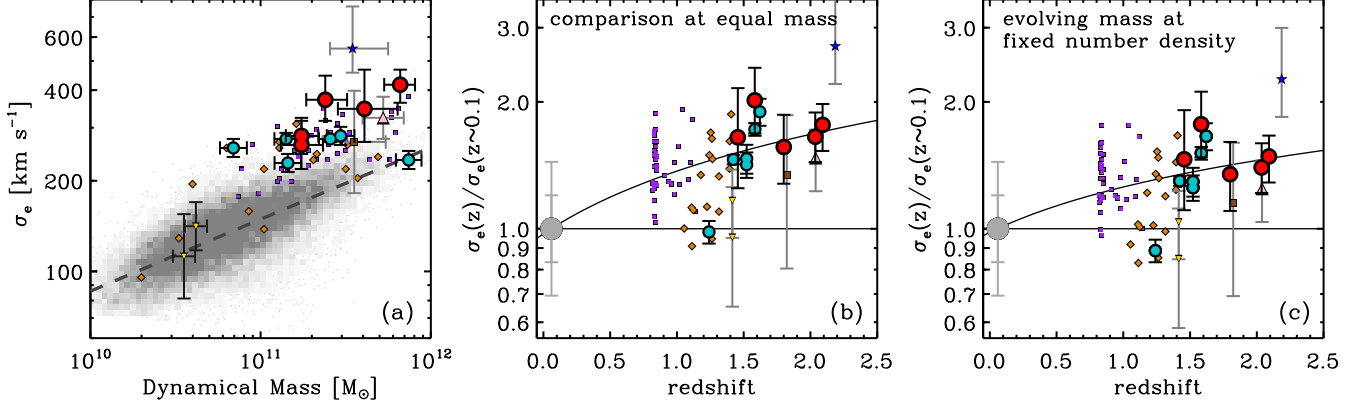


FIG. 9.— Redshift evolution in velocity dispersion. *Left panel:* σ_e versus dynamical mass. At fixed dynamical mass, we find that our galaxies have higher velocity dispersion as compared to low-redshift galaxies. *Middle panel:* Evolution of velocity dispersions at fixed dynamical mass versus redshift. We find the $\sigma_e \propto (1+z)^{0.50 \pm 0.08}$, i.e. at fixed dynamical mass the velocity dispersion decreases by a factor of ~ 1.7 from $z \sim 2$ to the present day. *Right panel:* Evolution of the velocity dispersion with an evolving mass function at constant number density. Opposite to the evolution in size, we now find a milder evolution of $\sigma_e \propto (1+z)^{0.36 \pm 0.08}$.

compared to the total luminosity. For a Sérsic profile this ratio is given by (Ciotti 1991):

$$\frac{L_{1\text{kpc}}}{L_{\text{tot}}} = \frac{\gamma(2n, x)}{\Gamma(2n)}. \quad (10)$$

Here, $\Gamma(2n)$ is the complete gamma function, $\gamma(2n, x)$ the incomplete gamma function, $x = b_n(r_{1\text{kpc}}/r_e)^{1/n}$, with $b_n = 1.9992n - 0.3271$. Using this ratio we can now calculate the dynamical mass within one 1 kpc and within r_e from the total mass:

$$M_{1\text{kpc}} = \frac{L_{1\text{kpc}}}{L_{\text{tot}}} M_{\text{dyn}}. \quad (11)$$

Here we assume that the dynamical mass profile follows the light profile, and furthermore that the mass-to-light ratio of the galaxy is radially constant. The detection of small color gradients in our galaxies indicates, however, that this is not the case, but the effect on the derived densities is small (Saracco et al. 2012; see also Szomoru et al. 2012). Finally, the densities are calculated as follows:

$$\rho_{1\text{kpc}} = \frac{4M_{1\text{kpc}}}{3\pi r_{1\text{kpc}}^3}, \quad (12)$$

and

$$\rho_e = \frac{2M_{\text{dyn}}}{3\pi r_e^3}. \quad (13)$$

Similarly, we now compare the density as a function of dynamical mass (see Figure 10). The top row shows the results for the density within one effective radius, while the bottom row compares the central density within 1 kpc. The first thing to notice is the large scatter for low-redshift galaxies when looking at ρ_e versus M_{dyn} , while $\rho_{1\text{kpc}}$ versus M_{dyn} shows a tight relation. The density-mass relation can be parametrized by:

$$\rho = \rho_c \left(\frac{M_{\text{dyn}}}{10^{11} M_\odot} \right)^b. \quad (14)$$

For the density within r_e we find $\rho_{c,e} = 4.7 \times 10^8 M_\odot \text{kpc}^{-3}$ and $b_e = -0.68 \pm 0.15$, and for the central density within 1 kpc $\rho_{c,1\text{kpc}} = 6.6 \times 10^9 M_\odot \text{kpc}^{-3}$ and $b_{1\text{kpc}} = 0.56$.

When comparing the galaxies in our high-redshift sample to galaxies in the SDSS, we find that they have higher densities within r_e . Comparison at equal dynamical mass shows that

the effective densities are higher by a factor of ~ 50 (Figure 10b) for our sample. The same comparison, but now for the central density within 1 kpc, reveals only mild evolution, approximately factor of ~ 3 from $z \sim 2$. When fitting

$$\rho \propto (1+z)^\alpha, \quad (15)$$

we find that $\alpha_e = 2.95 \pm 0.15$, while $\alpha_{1\text{kpc}} = 0.77 \pm 0.15$.

Instead of comparing galaxies at fixed mass, we again take into consideration that galaxies evolve in mass when comparing low- and high-redshift galaxies. This time, we find that ρ_e evolves even faster as compared to the equal mass comparison $\alpha_e = 3.35 \pm 0.15$. The density within 1 kpc, however, requires a decrease less than a factor of ~ 2 , $\alpha_{1\text{kpc}} = 0.44 \pm 0.15$, to evolve into present day galaxies.

7. DISCUSSION

In the previous section we have found that in order for the high-redshift galaxies in our sample to evolve into typical present-day galaxies, strong structural evolution is required. Effective radii need to increase, and the velocity dispersion within r_e has to decrease, while the density within the effective radius has to decrease by more than an order of magnitude. However, the central density can remain almost the same.

The dominant physical mechanism for this structural evolution is still a subject of ongoing debate. Size growth dominated by major mergers seems to be unlikely as it would increase the masses too much, which would make extremely massive galaxies too common in the local universe. As the mass and size increase at approximately at the same rate in major mergers, the galaxies would also remain too compact for their mass. Minor merging could offer a solution to the problem, as it can grow a galaxy in effective radius (r_e) steeper than $r_e \propto M_*$ (Naab et al. 2009; Bezanson et al. 2009; Hopkins et al. 2009b). In this scenario, the observed compact high-redshift galaxies may simply be the cores of local massive early-type galaxies, which grow inside-out by accreting (smaller) galaxies, and thus assemble a significant part of their mass at later times (van der Wel et al. 2009; Oser et al. 2010). In this section we will examine if dry minor merging agrees with our findings.

From a simple estimate, based on the virial theorem, Bezanson et al. (2009) predict how the effective radii changes if a massive galaxy undergoes a series of minor mergers. With

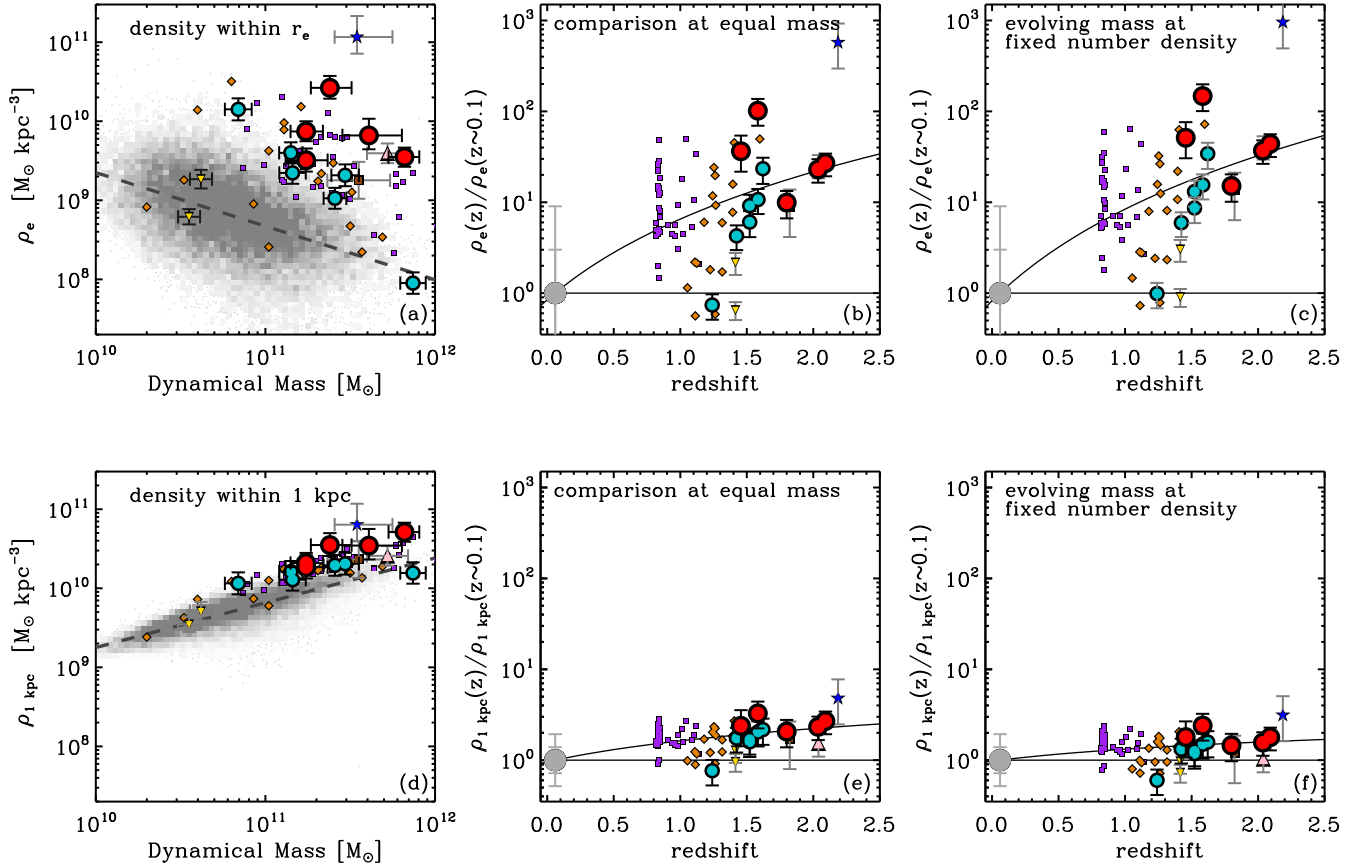


FIG. 10.— Redshift evolution of the central and effective densities, as calculated according to Equation 12 & 13. Top row shows the results for the effective mass density, while the bottom row shows the mass density within 1 kpc. *Left panels:* The density within r_e versus dynamical mass (*top*), we find that low-redshift galaxies show a large scatter at fixed dynamical mass. Still, our high-redshift galaxies have higher densities within r_e at fixed mass. In contrast to ρ_e , the density within 1 kpc versus dynamical mass (*bottom*) shows very little scatter, and we find only a small difference between low- and high-redshift galaxies. *Middle panels:* Evolution in the density at fixed dynamical mass versus redshift. We find a strong evolution for the effective density (*top*) with $\rho_e \propto (1+z)^{2.95 \pm 0.15}$, or a decrease by a factor of ~ 2 from high-redshift to $z \sim 0$. For the central density, however, we only find that $\rho_{1kpc} \propto (1+z)^{0.77 \pm 0.15}$, or a decrease of ~ 2.3 . *Right panels:* Evolution of the density with an evolving mass function at constant number density. Similar to the effective radii, we find a stronger evolution for the effective density, i.e. $\rho_e \propto (1+z)^{3.35 \pm 0.15}$, while the central density requires very little to no evolution $\rho_{1kpc} \propto (1+z)^{0.44 \pm 0.15}$.

only eight 1:10 mergers, the effective radii can grow by a factor of ~ 5 while only having the mass increase by a factor of ~ 2 . This is also described by Naab et al. (2009), who state that if an initial system undergoes a mass increase by a factor of 2 due to accretion of very small systems, then the final radius of the system is four times larger, the velocity dispersion is reduced by a factor of 2, and the density is reduced by a factor of 32. This prediction is confirmed from by their hydrodynamical cosmological simulation which gives similar size evolution as presented here.

Using hydrodynamic simulations of galaxy mergers, Hopkins et al. (2009a), also find evidence for size evolution. When comparing effective radii at fixed mass, they find an evolution in size of $r_e \propto (1+z)^{-0.48}$ for galaxies with $\log M_*/M_\odot = 11$, which is weaker than found by this study and many others. Oser et al. (2012) find a size evolution in their hydro simulation, which is much stronger: $r_e \propto (1+z)^{-1.44}$, on the high side of current results.

Oser et al. (2012) find a similar evolution in velocity dispersion of $\sigma \propto (1+z)^{0.44}$, to that found in this work. In contrast, Hopkins et al. (2009a), predict such systems have roughly the same or at most a factor ~ 1.25 larger velocity dispersions.

Evolution of the density is also discussed in both Bezanson et al. (2009) and Hopkins et al. (2009b). Based on photometric data, both studies find that while the den-

sity within one effective radius is higher at high-redshift, the central density of high-redshift galaxies is very similar to local massive ellipticals. From hydro simulations, Naab et al. (2009) show that the central density within 1 kpc decreases by a factor of 1.5 from $z = 2$ to $z = 0$, caused by dynamical friction from the surviving cores. Similar, Oser et al. (2012) show that the central density evolves only weakly, while the density within r_e decreases rapidly by more than an order of magnitude, in good agreement with what we find here.

This is in contrast with the findings of Saracco et al. (2012), who find no evidence for higher effective densities when comparing their $z \sim 1.5$ galaxies to low-redshift cluster galaxies. Furthermore, the large scatter that they observed in the effective density and the apparent evolution, is simply due a peculiar feature in the Sérsic profile.

In Figure 6a we showed that the ratio of M_*/M_{dyn} may have evolved from $z \sim 2$ to $z \sim 0$. As compared to SDSS galaxies with $\log M_*/M_\odot > 11$, we find that the median M_*/M_{dyn} is higher by 50% at $z > 1.5$, and that $M_*/M_{dyn} \propto (1+z)^{0.17 \pm 0.04}$. However, this result is still uncertain due to the selection effects inherent in this sample and large measurement errors on both masses. We note that this effect is predicted by simulations; as the effective radius of a galaxy evolves, the dark matter fraction within r_e will also increase. Hopkins et al. (2009a) predicts evolution by a factor of ~ 1.25 for galaxies

with $\log M_*/M_\odot = 11$, with the effect increasing with stellar mass. Hilz et al. (2012) also find a strong evolution in the dark matter fraction in their hydro-simulation, and predict that quiescent galaxies at $z \sim 2$ have lower dark matter fractions ($\gtrsim 80\%$). They mention that is mainly driven by the strong size increase, which therefore probes a larger region that is dominated by dark matter.

8. SUMMARY AND CONCLUSION

In this paper, we present deep UV-NIR spectroscopy of 5 massive ($> 10^{11} M_\odot$) galaxies at $z \sim 2$, using X-Shooter on the VLT. These spectra enable us to measure stellar velocity dispersions with higher accuracy than done before at this redshift: we triple the sample of $z > 1.5$ galaxies with well constrained ($\delta\sigma < 100 \text{ km s}^{-1}$) velocity dispersion measurements. We find that the stellar velocity dispersions are high ($290\text{--}450 \text{ km s}^{-1}$) compared to equal-mass galaxies in the SDSS.

We combine these kinematic results with size measurements using GALFIT on HST-WFC3 H_{160} and UDS K-band imaging, and use these measurements to derive dynamical masses. Stellar masses are obtained from stellar population synthesis modeling on the VIS-NIR spectra in combination with the available broad- and medium-band data. The SPS-modeling shows that our galaxies have ages ranging from 0.5–2 Gyr, and show no signs of on-going star formation. We complement our results with stellar kinematic results from other studies at low and high redshift to study the structural evolution of massive quiescent galaxies.

We find good correspondence between the dynamical and stellar masses, with the dynamical mass being higher by $\sim 15\%$. Our results suggest that stellar mass measurements for quiescent galaxies at high-redshift are robust.

At fixed dynamical mass we find that the effective radius increases by a factor of ~ 2.8 , while the velocity dispersion decreases by a factor of ~ 1.7 from $z \sim 2$ to the present day. Furthermore, we study how the mass density within r_e and 1 kpc evolves with time. While we find a strong decrease of the mass density within one effective radius (factor of ~ 21), it only decreases mildly within 1 kpc (factor of ~ 2.3). Instead of comparing galaxies at fixed dynamical mass, we also use an evolving mass limit as defined by fixed number density. By accounting for concurrent mass growth in our comparison of high- and low-redshift galaxy populations, we find an even stronger evolution in galaxy sizes (factor of ~ 4). We find that velocity dispersion decreases less dramatically with time, differing by only a factor of ~ 1.4 . Finally, for the mass density within r_e , we find a stronger evolution, but interestingly, the mass density within 1 kpc is consistent with no evolution.

We examine if our results are compatible with the current idea that massive quiescent galaxies at $z \sim 2$ grow inside out through dry minor mergers. From hydrodynamical simulations, Oser et al. (2012) find similar results to those shown here. Even though their size evolution is somewhat stronger, they do find a similar decrease in the velocity dispersion. Furthermore, their simulations show that the central density re-

mains almost the same, while at the same time the effective density decreases by more than an order of magnitude, in good agreement with our findings.

Finally, we find that even though the stellar masses are consistent with the dynamical masses, the ratio of M_*/M_{dyn} may decrease with time. This, too, is predicted by minor merging simulations, which show that the size growth due to minor merging is accompanied by an increase of the dark matter fraction within an effective radius. This is due to the fact that the dark matter profile is less steep than the stellar mass profile, and thus the dark matter to stellar mass fraction increases with radius.

Despite the vastly improved accuracy of our derived dynamical masses and stellar population parameters, the broader inferences of our study is still limited by the very small number of high-redshift galaxies with such information. We have shown that our sample is biased towards younger galaxies, compared to a stellar mass limited sample at $z \sim 2$, with smaller effective radii as compared to the full population of quiescent galaxies $z \sim 2$. Only with a larger unbiased sample of massive quiescent galaxies at high redshift can we start to comprehend the final phase that massive galaxies go through in becoming today's ellipticals.

We thank Daniel Szomoru for providing his residual-correct code and the galaxy sizes for GOODS-S CANDELS; Andrew Newman for providing the corrected stellar masses; Shannon Patel for the structural parameters of galaxies in UDS CANDELS; and Ivo Labbé for helpful comments regarding the SFRs from 24 micron flux. It is a pleasure to acknowledge the contribution to this work by the NMBS collaboration. We would also like to thank Adam Muzzin for useful discussions. This research was supported by grants from the Netherlands Foundation for Research (NWO), the Leids Kerkhoven-Bosscha Fonds. Support for program HST-GO-12167.1 was provided by NASA through a grant from the Space Telescope Science Institute.

This work is based on observations taken by the CANDELS Multi-Cycle Treasury Program with the NASA/ESA HST, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

This publication also makes use of the Sloan Digital Sky Survey (SDSS). Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>. The SDSS Participating Institutions are the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Max Planck Institut fur Astronomie, the Max Planck Institut fur Astrophysik, New Mexico State University, Princeton University, the United States Naval Observatory, and the University of Washington.

APPENDIX

ROBUSTNESS OF THE VELOCITY DISPERSION MEASUREMENTS

As it has only recently become possible to measure velocity dispersions at high redshift, the stability of these measurements has barely been tested. In this appendix, we will study the effect of fitted wavelength range, template choice, degree of the additive

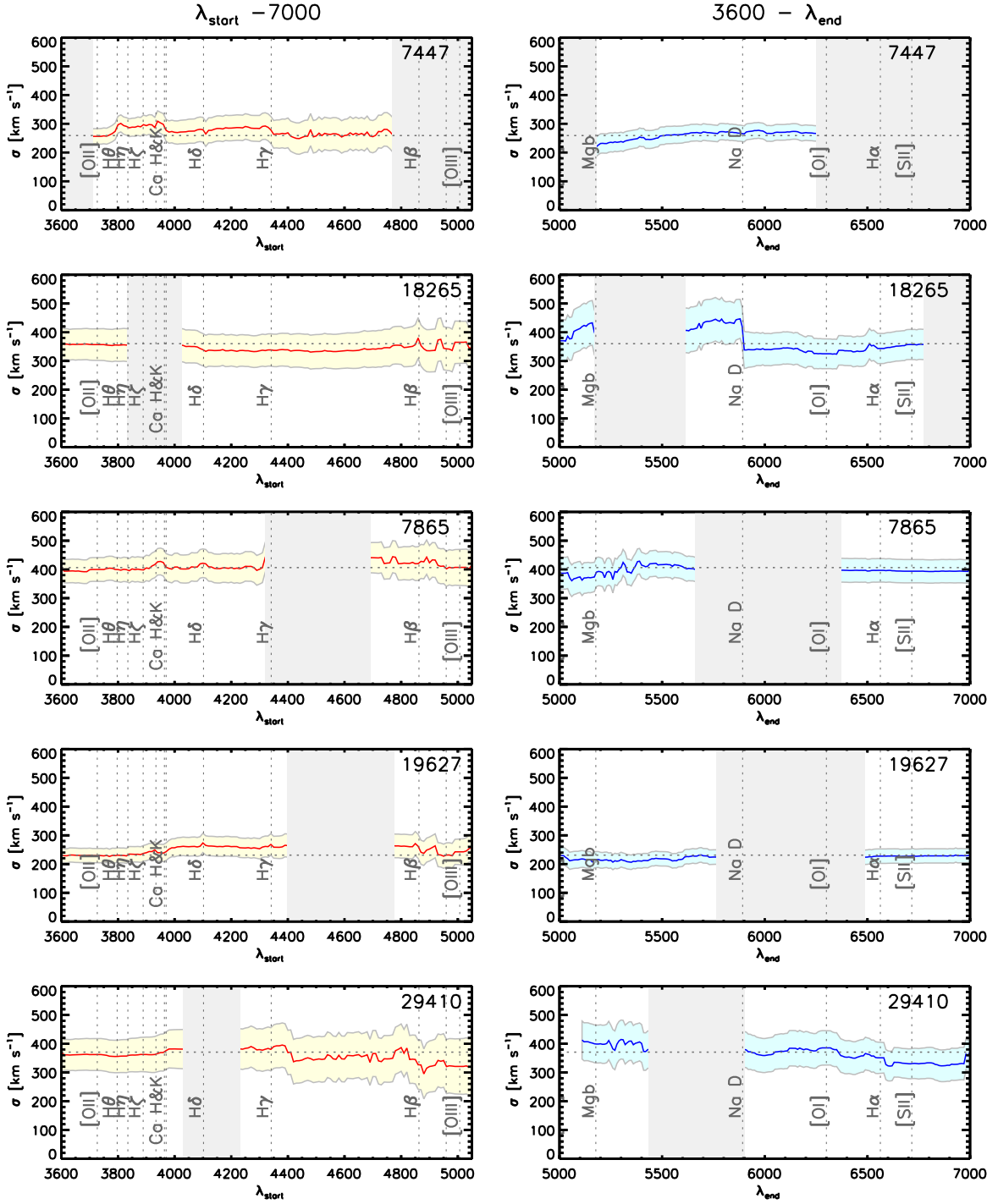


FIG. 11.— Velocity dispersion versus wavelength range. Left panel shows the results for $\lambda_{\text{start}} < \lambda < 7000\text{\AA}$, while the right panel is for $3600\text{\AA} < \lambda < \lambda_{\text{end}}$. The horizontal shaded region indicates the $1\text{-}\sigma$ error from pPXF, the dashed horizontal line indicates the velocity dispersion when fitting the full range. Most prominent absorption lines are indicated, and the regions affected by strong sky-lines and atmospheric absorption are shown in gray. Overall, we find a stable solution for the velocity dispersion while changing the wavelength range. For NMBS-C7447 and UDS-19627 we do find a small increase in the region around the Balmer break. For NMBS-C17637, we find that excluding Na D has a great impact on the velocity dispersion.

polynomial, and choice of stellar populations models.

Dependence of the Velocity Dispersion on the Wavelength Range

Our sample has a redshift range of $1.4 < z < 2.1$, which means that different parts of the rest-frame spectra will be affected by sky-lines and atmospheric absorption for each galaxy. This can be seen from Figure 4, which shows that we often lose strong absorption features in our spectrum, which affects the region of the rest-frame spectrum that we can use for the fit. Here we investigate how stable the measured velocity dispersion is as a function of the wavelength range.

For the velocity dispersion fitting in this Paper, the lower wavelength limit is set by stellar libraries and models, as no systematic

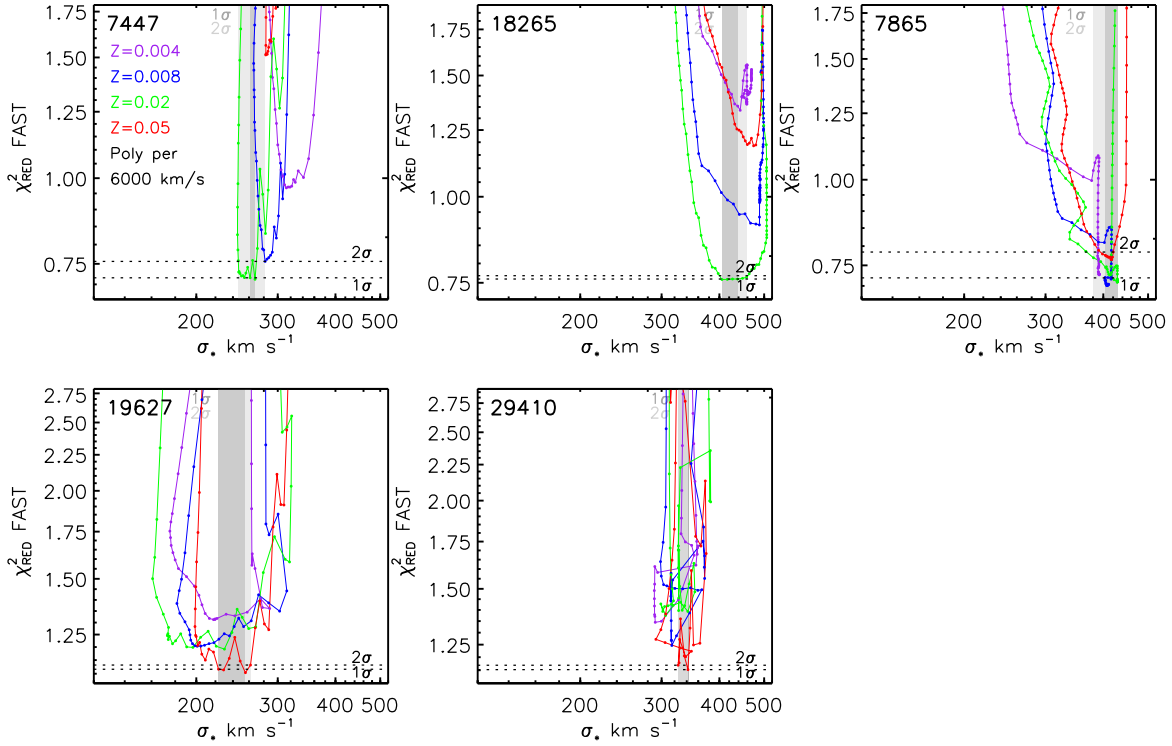


FIG. 12.— Reduced χ^2 from SPS-modeling versus velocity dispersion. Different points indicate different template ages, while different colors show different metallicities. All fits are done on the full-wavelength range and with one-order per 6000 km s⁻¹ for the additive polynomial. We find a stable solution for the velocity dispersion by using different templates, except for UDS-19627, which shows a large range in velocity dispersions.

high-resolution observations exist below 3550Å. The higher wavelength limit is set by lower S/N in our spectra in the observed K-Band. Our approach for testing the wavelength dependence of the fit is as follows. First, we use the full-range spectrum to determine a best-fit polynomial (1 order per 10 000 km s⁻¹), which is used to correct for the difference between the observed and the template continuum. Next, we repeat the measurement with a zeroth-order polynomial while changing the start or end wavelength. The polynomial is not a free parameter in this fit, as this would make it impossible to separate between the effect of the polynomial and the wavelength range. Note also, that we use a single template for all fits as determined from the full Visual+NIR spectrum together with broad- and medium-band data (see Section 3.2).

Figure 11 shows the results for the different sources. The left panel shows the result where we change the starting wavelength, i.e. the wavelength range is $\lambda_{start} < \lambda < 7000\text{\AA}$, whereas the panels on the right show the effect of changing the end wavelength $3600\text{\AA} < \lambda < \lambda_{end}$. The first thing to notice is that the measured dispersions are remarkably stable, even when most of the absorption lines have been excluded from the fit. The two galaxies with the youngest ages and strong Balmer absorption lines (NMBS-C7447 and UDS-19627) show a change in the velocity dispersion in the region of the Balmer break. NMBS-C7447 shows an increase for $\lambda_{start} > 3800\text{\AA}$, but decreases after Ca H&K have been removed from the fit. UDS-19627 shows an increase of 50 km s⁻¹ when the Balmer break is excluded. When we reduce the red part of the spectrum (Figure 11, right panels), we also find a very stable fit, except for NMBS-C17637. After excluding Na D from the fit, the velocity dispersion increases by a ~ 100 km s⁻¹. We think this is because NMBS-C17637 has a more evolved stellar populations than say for example NMBS-C7447. With the Ca H&K lines masked out due to atmospheric absorption, Na D is one of the strongest lines in the spectrum, and its exclusion could explain the sudden change in the measured dispersion.

To summarize, for most galaxies we find no strong dependence on the wavelength range that is used in the fit. We do find that with decreasing wavelength range, the random error increases. Finally, even in the absence of strong absorption features like Ca H&K, we find similar velocity dispersions as compared to the full range fit.

Dependence of the Velocity Dispersion on the Template Choice

Next, we study how different templates may influence the measured velocity dispersion. We use a sample of BC03 τ -models, as presented in Section 3.2. In particular, we are interested in the effect of template age and metallicity. In Figure 12 we show the reduced χ^2 from the SPS-modeling, versus the velocity dispersions measured using this template. The different points represent different ages of the templates, with the minimum χ^2 corresponding to the best-fit age as listed in Table 2. Different colors indicate the different metallicities. We show the χ^2 from the SPS-modeling instead of the dispersion fit, as the former is derived from the full Visual and NIR spectrum plus all the broad- and medium-band photometric data. This yields better constraints for the stellar population, thus larger relative ranges in χ^2 as compared to the χ^2 from the dispersion fit. Also, as we add high-order additive polynomials to the templates before fitting the velocity dispersion (in this case 1 order per 6 000 km s⁻¹), the effect of different template-ages is mostly washed out, and we get a small relative χ^2 .

In Figure 12 we see that for most galaxies the velocity dispersion for templates allowed within 2σ give consistent results.

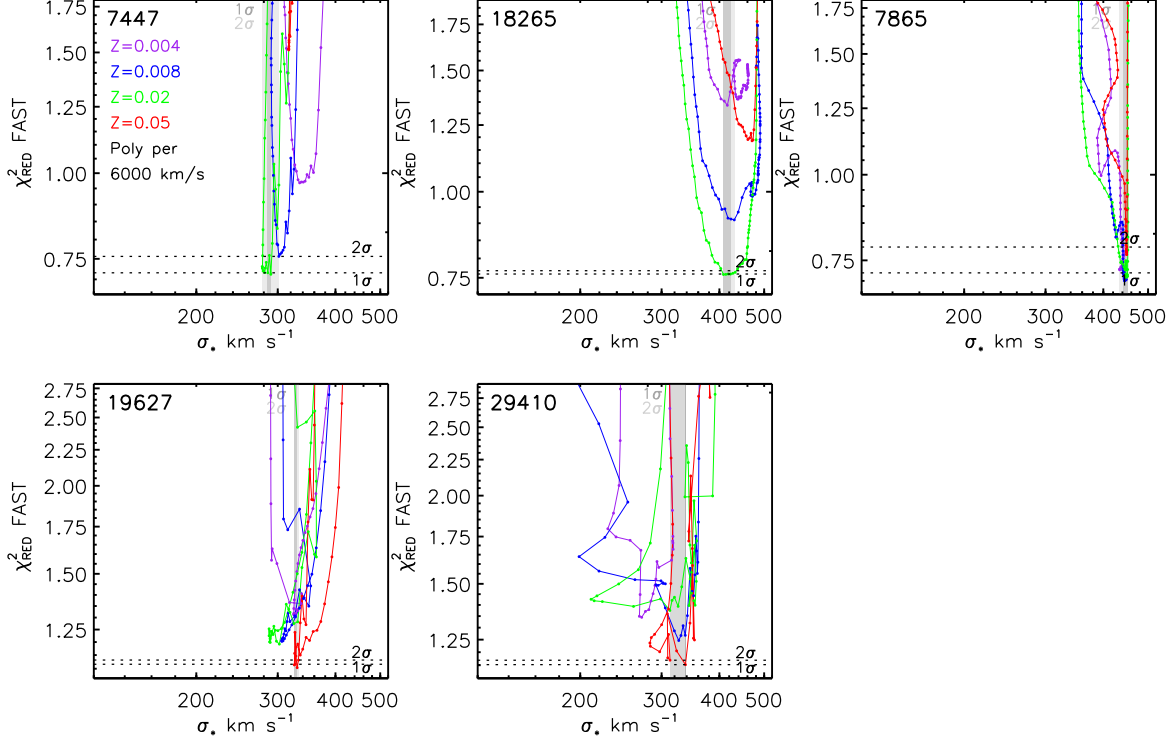


FIG. 13.— Reduced χ^2 versus velocity dispersion, similar to Figure 12 but now excluding the Balmer break ($\lambda > 4020 \text{ \AA}$). Whereas with the full wavelength range UDS 19627 showed a large variation in the velocity dispersion as a function of template, this time we do find a stable solution. The other sources also, show a decrease in the uncertainty due to different templates, except for UDS-29410. However, this is due to lower S/N of this spectrum and the shorter fitted wavelength range as before.

Different metallicities do give different velocity dispersions at their minimum χ^2 , but this is a mere reflection of the age-metallicity degeneracy. Different metallicities have different best-fitting ages, which in turn give different velocity dispersion. At a $1 - \sigma$ level, we only have a handful of best-fitting templates, from which we obtain similar velocity dispersions. UDS-19627 is the exception, which shows a large dependency of the measured velocity dispersion as a function of template age. At the $1 - \sigma$ level, we find a range of $\sim 30 \text{ km s}^{-1}$, due to template uncertainty, while the random error is one of the lowest, only $\sim 30 \text{ km s}^{-1}$. Even though templates with χ^2 of 1.20 or higher are statistically considered a bad fit, the large range in the velocity dispersion is worrying. If we could not have constrained the best-fitting template from the full range spectrum and broad band data, the measured velocity dispersion would be highly uncertain.

However, if we exclude the Balmer break from the velocity dispersion fit, the dependency on template age almost completely disappears (Figure 13). The case of UDS-19627 is the most dramatic one, where we suddenly see a tight range of best-fitting velocity dispersions. We do note that the velocity dispersion has increased, as was already shown in Figure 11. UDS-29410 appears to have a slightly higher template uncertainty if we only fit for $\lambda > 4020 \text{ \AA}$, but this is driven by the lower S/N of the spectrum as we now use a shorter wavelength range.

To conclude, we find a systematic uncertainty due to templates with different ages. This is caused by the Balmer break, present in the relatively young galaxies in our sample. By only fitting for $\lambda > 4020 \text{ \AA}$, the uncertainty due template mismatch almost completely disappears. In that case, templates with different metallicities do give different velocity dispersions, but this is most likely caused by the underlying age-metallicity degeneracy.

Dependence of the Velocity Dispersion on the Order of the Additive Polynomial

In order to correct for stellar continuum emission differences between the observed galaxy spectrum and the template, we use an additive Legendre polynomial. If we would not apply such a correction, the fitting routine could try to correct for this discrepancy by changing the velocity dispersion. Values that are typically used in the literature vary from 5000 to 15000 km s^{-1} per order. We do not use multiplicative Legendre polynomials, because the S/N of the spectra are too low, and it would add another degree of uncertainty to the fit. Here, we test the influence of the additive polynomial to our measured velocity dispersions. Again we use the best-fit τ -model as a template while varying the additive polynomials from 0 to 50. We fit both the full-wavelength range and the wavelength range with $\lambda > 4020 \text{ \AA}$.

Figure 14 shows the results, with the blue line representing the full-wavelength fit, while the red line shows the results for $\lambda > 4020 \text{ \AA}$. The vertical dashed line indicates the polynomial with one order per 10000 km s^{-1} . Overall we find that by increasing the additive polynomial, the measured velocity dispersion increases. In general we find that between a polynomial of 10 and 30, the smallest increase in the velocity dispersion occurs, and this appears to be the most stable region. For this reason we use one order per 10000 km s^{-1} for our science results.

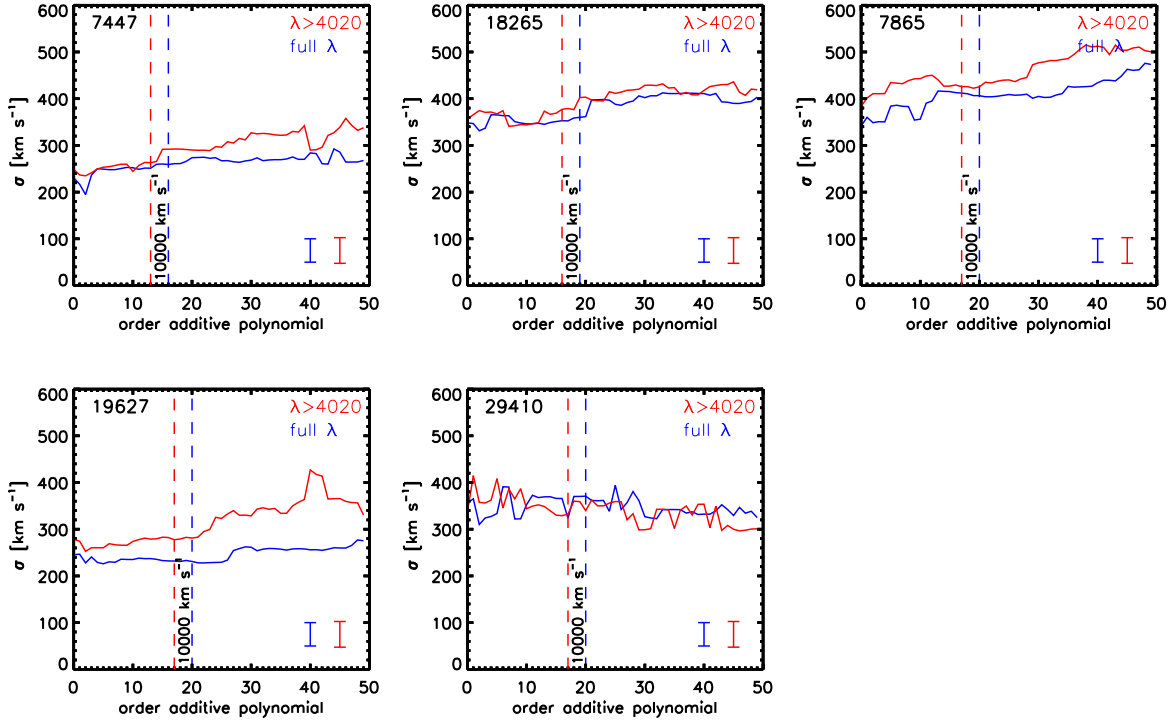


FIG. 14.— Velocity dispersion as measured with different order additive polynomials. The full-wavelength range fit is shown in blue, while the fit with $\lambda > 4020 \text{ \AA}$ is shown in red. We find an increase of the velocity dispersion with increasing order. The fit is most stable when using a polynomial with order between 10 and 30. The vertical dashed lines shows the order that corresponds to 10000 km s^{-1} , the polynomial that we use for obtaining science results.

Dependence of the Velocity Dispersion on the SPS-models and Template Construction

Here, we test how our velocity dispersions would change if we make different choices for the SPS-model, and test the difference between a single τ -model and a template constructed from different SSP models.

The left and middle panel of Figure 15 show what would happen if we would choose the Flexible-SPS models by Conroy et al. (2009) (CO9) or the models by Maraston & Strömbäck (2011) (MA11). These models are based on a different stellar library with slightly higher resolution as compared to BC03, i.e. MILES (Sánchez-Blázquez et al. 2006) versus STELIB (Le Borgne et al. 2003). If a systematic uncertainty in the measured velocity dispersion is present, for example due to the resolution or details that go into the SPS-models, it would show up in this comparison. We determined a best-fit τ model using the CO9 and MA11 models in exactly the same way as was done for the BC03 models (See Section 3.2). When comparing the velocity dispersions derived by using the BC03 and CO9 τ -models, we do not find any significant difference (right panel of Figure 15). In the middle panel we compare MA11 and BC03. Here too, we find a good correspondence, but for UDS19627 the MA11 models give a lower velocity dispersion, but still within the $1 - \sigma$ error. These results confirms that our measurements are stable against different SPS-models and templates with different resolution.

Finally, in Figure 15 (right panel) we compare the velocity dispersion from the best-fit τ -model, versus an optimal template constructed by pPXF. This optimal template is build from a full spectral library from BC03 Single Stellar Population (SSP) models, with full range in age and metallicity. We note that the optimal template construction only uses the wavelength regime provided in the dispersion fit ($3600 \text{ \AA} < \lambda < 7000 \text{ \AA}$), and does not take into account the effects of dust. The velocity dispersions from the optimal templates are slightly higher as compared to the single τ -model, although it is well below the random error. Interestingly, the galaxy with the largest dust-contribution, NMBS-C17637, also shows the largest difference between the two different fitting techniques. In this paper, we choose to use the best-fitting τ -model as the dispersion template, as this is the best representation of the stellar population. As the stellar mass is also based on this τ -model, we use the same stellar population when comparing stellar to the dynamical mass.

APERTURE CORRECTIONS FOR VELOCITY DISPERSION MEASUREMENTS

Here, we investigate the effects of different apertures and extraction methods on the observed velocity dispersion. The standard approach is to use a power-law to scale the observed velocity dispersion, measured within a certain r_{aper} , to the velocity dispersion as if measured within $r_{e/8}$, using the following expression by Jorgensen et al. (1995):

$$\frac{\sigma_{\text{ap}}}{\sigma_{e/8}} = \left(\frac{r_{\text{aper}}}{r_{e/8}} \right)^{-0.04}. \quad (\text{B1})$$

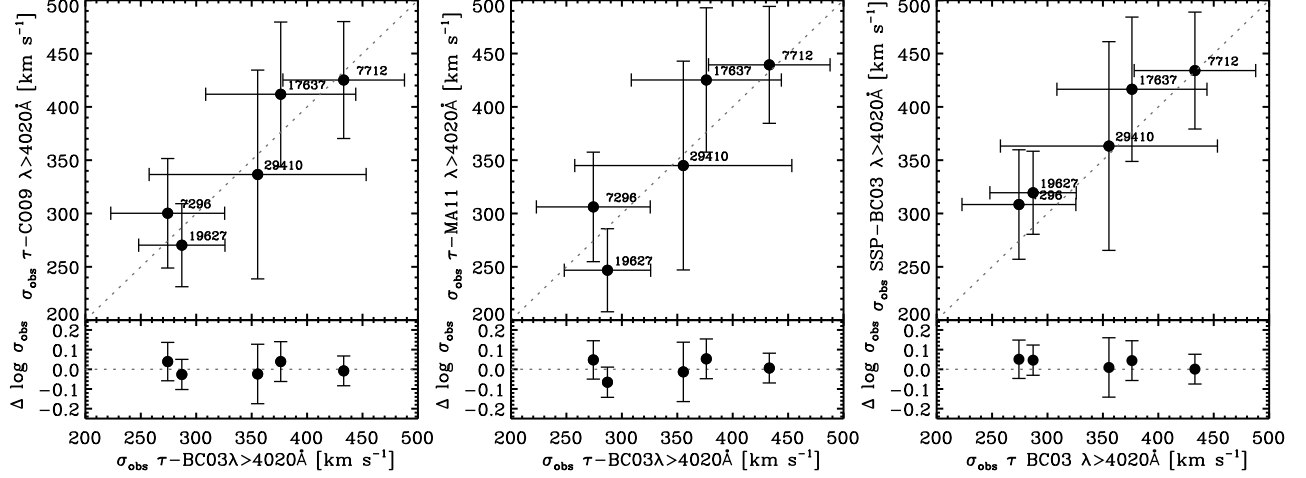


FIG. 15.— Comparison of the velocity dispersion when using different methods. *Left panel:* Velocity dispersion as derived with the models of BC03 and Conroy et al. (2009). Even though the models are based on different stellar libraries with different resolution, the velocity dispersion measured with these models are consistent. *Middle panel:* Similar as the left panel but now using BC03 and Maraston & Strömbäck (2011) for the comparison. In general we find consistent results within the errors, but UDS19627 has a lower velocity dispersion when we use the MA11 models. *Right panel:* Velocity dispersion as derived with optimal template construction versus the best-fitting τ -model. The optimal template is constructed from BC03 SSP models with a full range in age and metallicity. Dispersions from the optimal template construction are slightly higher, but well within the errors.

Or using a more recent result from the SAURON survey (Cappellari et al. 2006):

$$\frac{\sigma_{\text{ap}}}{\sigma_e} = \left(\frac{r_{\text{aper}}}{r_e} \right)^{-0.066 \pm 0.035} \quad (\text{B2})$$

In the case of Jorgensen et al. (1995), the velocity dispersions have been measured in the range of $-1.5 < \log(r_{\text{aper}}/r_e) < -0.5$, and for Cappellari et al. 2006 for $-1.0 < \log(r_{\text{aper}}/r_e) < 0$. In this work however, we are outside this range with $\log(r_{\text{aper}}/r_e) > 0.5$, that is r_e being much smaller than r_{aper} . Also, the galaxies in our sample have effective radii much smaller than the FWHM of the PSF, which is why the standard approach most likely will not be valid.

We analyze this problem in two steps. Firstly, we use an analytic description for the kinematic profile and match this to the observed relation as found by Cappellari et al. (2006). Secondly, using our model we study the behavior of the observed velocity dispersion for different apertures (both circular and non-circular) and different FWHM for the PSF. Our reference model will be a circular aperture with size r_e , without the effects of seeing.

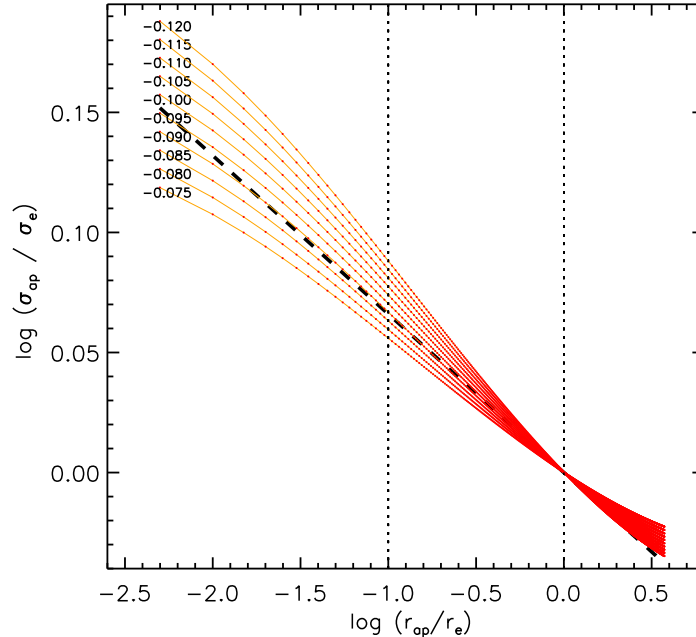


FIG. 16.— Observed kinematic profile in a circular aperture. Best fit is $d = -0.089$

From Figure 18 it is clear, that by using a simple power law for the aperture correction, we would overestimate the corrections by a large fraction. Furthermore, we have shown that it is vital to use a suitable aperture and include the effects of seeing, especially when $r_{\text{aper}}/r_e > 2$.

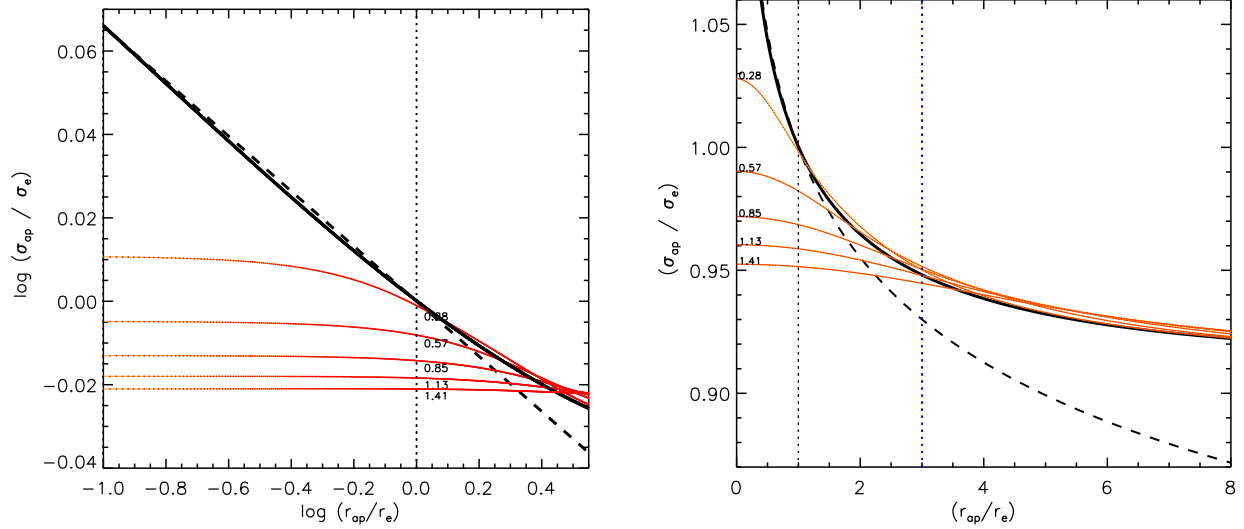


FIG. 17.— Observed kinematic profile in a circular aperture, while varying the FWHM of the psf (red) and the Cappellari et al. relation (dashed). The black line shows our reference model, i.e. circular aperture with no influence of the PSF.

A good description of the kinematic profile for early galaxies is given by (see e.g. True et al. 1999, Bertin et al 2002) :

$$\sigma(r) = \left(\frac{r}{r_e}\right)^d \times \sigma_{cst}, \quad (\text{B3})$$

with $-0.1 < d < 0$. The observed kinematic profile within a circular aperture will be a projection of $\sigma(r)^2$ and the galaxies intensity profile :

$$\sigma^2(r_{aper}) = \frac{\int_0^{r_{aper}} \sigma^2(r) I_{gal}(r) 2\pi r dr}{\int_0^{r_{aper}} I_{gal}(r) 2\pi r dr}. \quad (\text{B4})$$

Here, $I_{gal}(r)$ is the Sérsic Profile:

$$I(r) = I_e \exp \left\{ -b_n \left[\left(\frac{r}{r_e} \right)^{1/n} - 1 \right] \right\} \quad (\text{B5})$$

To avoid numerical issues in the center, we approximate eq. B3 and eq. B5 by:

$$\sigma(r) = \left(\frac{r + r_{core}}{r_e} \right)^d \times \sigma_{cst}, \quad (\text{B6})$$

$$I(r) = I_e \exp \left\{ -b_n \left[\left(\frac{r + r_{core}}{r_e} \right)^{1/n} - 1 \right] \right\} \quad (\text{B7})$$

where r_{core} is chosen to be $1/30 r_e$. We can estimate the power d by evaluating Eq. B4 for different values of d and comparing the results to Eq. 1 from Cappellari et al. (2006).

Figure 16 shows the Cappellari et al. relation (dashed) and our model with different values for d , normalized to r_e . We find a best fitting value for $d = -0.089$ in the region of $(-1.0 < \log(r_{aper}/r_e) < 0)$. Notice that for $\log(r_{aper}/r_e) > 0.0$ our model deviates from the simple power law in Eq. B3.

Now that we have found the intrinsic kinematic profile, we can explore the influence of the PSF on the observed dispersion. The PSF is modelled using a combination of two gaussians, where $\sigma_1 = 2\sigma_2$, and both gaussians having equal flux (reference). This PSF is then convolved with the kinematic and intensity profiles:

$$\sigma^2(r_{aper}) = \frac{\int_0^{r_{aper}} \left\{ \left[\sigma^2(r) I_{gal}(r) \right] \otimes PSF \right\} 2\pi r dr}{\int_0^{r_{aper}} \left\{ \left[I_{gal}(r) \right] \otimes PSF \right\} 2\pi r dr}. \quad (\text{B8})$$

The results for PSFs with different values of σ are shown in Figure 17

Instead of using a circular aperture, we now consider the case of a rectangular aperture, similar to what is used in the spectra from X-Shooter. Eq. B8 can be modified to include a weight function $g(y)$ which is commonly used in optimized extraction. The integral in Eq. B9 is also replaced by a Riemann sum. In the x direction (slit width) the aperture size is always the same, i.e. $0.9''$ for X-Shooter spectra, but the y direction is now parallel to r_{aper} .

$$\sigma^2(x_{aper}, y_{aper}) = \frac{\sum_0^{x_{aper}} \sum_0^{y_{aper}} \left\{ \left[\sigma^2(x, y) I_{gal}(x, y) \right] \otimes PSF \right\} g(y) \Delta x \Delta y}{\sum_0^{x_{aper}} \sum_0^{y_{aper}} \left\{ \left[I_{gal}(x, y) \right] \otimes PSF \right\} g(y) \Delta x \Delta y}. \quad (B9)$$

Figure 18 shows the difference between using a circular, rectangular, and rectangular aperture including a weighting function. The slit width that was used is $0.9''$, with the FWHM of the PSF also being $0.9''$. The spectrum was extracted with $r_{aper} = 0.45''$. We see that the correction is slightly higher for the rectangular aperture as compared to the circular aperture at $r_{aper}/r_e = 2.25$. The behavior at $r_{aper}/r_e > 3$ is very different for the three different cases. When using a rectangular aperture with optimized extraction, the observed profile is flatter as compared to the other models. The corrections are on average between 3 and 5%.

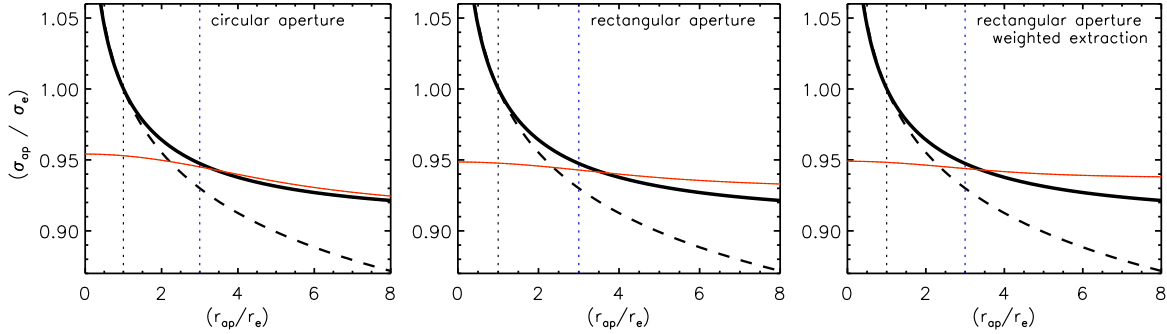


FIG. 18.— Observed kinematic profile in different apertures: circular, rectangular, and rectangular including a weight function. Dashed line is the Cappellari et al. relation, solid line is our reference model (circular aperture, no PSF), and the red line is the modeled kinematic including the effect of the PSF. The vertical blue dashed line shows a typical r_{aper}/r_e for our observations. We note that r_{aper} is in the direction along the slit.

REFERENCES

- Bezanson, R., van Dokkum, P. G., Tal, T., Marchesini, D., Kriek, M., Franx, M., & Coppi, P. 2009, *ApJ*, 697, 1290
- Bezanson, R., van Dokkum, P., van de Sande, J., Franx, M., & Kriek, M. 2012, arXiv:1210.7236
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- Bertin, G., Ciotti, L., & Del Principe, M. 2002, *A&A*, 386, 149
- Blakeslee, J. P., Holden, B. P., Franx, M., et al. 2006, *ApJ*, 644, 30
- Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, *AJ*, 129, 2562
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2010, *ApJ*, 709, L133
- Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2011, *ApJ*, 739, 24
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, 351, 1151
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, *ApJ*, 533, 682
- Cappellari, M., & Emsellem, E. 2004, *PASP*, 116, 138
- Cappellari, M., et al. 2006, *MNRAS*, 366, 1126
- Cappellari, M., et al. 2009, *ApJ*, 704, L34
- Cassata, P., et al. 2010, *ApJ*, 714, L79
- Cenarro, A. J., & Trujillo, I. 2009, *ApJ*, 696, L43
- Chabrier, G. 2003, *PASP*, 115, 763
- Ciotti, L. 1991, *A&A*, 249, 99
- Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, 699, 486
- Daddi, E., et al. 2005, *ApJ*, 626, 680
- D’Odorico, S., et al. 2006, *Proc. SPIE*, 6269,
- Ferré-Mateu, A., Vazdekis, A., Trujillo, I., et al. 2012, *MNRAS*, 423, 632
- Förster Schreiber, N. M., Franx, M., Labbé, I., et al. 2006, *AJ*, 131, 1891
- Franx, M., van Dokkum, P. G., Schreiber, N. M. F., Wuyts, S., Labbé, I., & Toft, S. 2008, *ApJ*, 688, 770
- Goldoni, P., Royer, F., François, P., Horrobin, M., Blanc, G., Vernet, J., Modigliani, A., & Larsen, J. 2006, *Proc. SPIE*, 6269,
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35
- Hilz, M., Naab, T., & Ostriker, J. P. 2012, arXiv:1206.5004
- Hopkins, P. F., Hernquist, L., Cox, T. J., Keres, D., & Wuyts, S. 2009a, *ApJ*, 691, 1424
- Hopkins, P. F., Bundy, K., Murray, N., et al. 2009b, *MNRAS*, 398, 898
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33
- Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, *ApJS*, 172, 196
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36
- Kormendy, J., Fisher, D. B., Cornell, M. E., & Bender, R. 2009, *ApJS*, 182, 216
- Kriek, M., et al. 2006, *ApJ*, 649, L71
- Kriek, M., van Dokkum, P. G., Labbé, I., Franx, M., Illingworth, G. D., Marchesini, D., & Quadri, R. F. 2009, *ApJ*, 700, 221
- Labbé, I., et al. 2005, *ApJ*, 624, L81
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, 379, 1599
- Le Borgne, J.-F., Bruzual, G., Pelló, R., et al. 2003, *A&A*, 402, 433
- Maraston, C., & Strömbäck, G. 2011, *MNRAS*, 418, 2785
- Massey, R., Stoughton, C., Leauthaud, A., et al. 2010, *MNRAS*, 401, 371
- Munari, U., Sordo, R., Castelli, F., & Zwitter, T. 2005, *A&A*, 442, 1127
- Muzzin, A., Marchesini, D., van Dokkum, P. G., et al. 2009, *ApJ*, 701, 1839
- Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, *ApJ*, 699, L178
- Newman, A. B., Ellis, R. S., Treu, T., & Bundy, K. 2010, *ApJ*, 717, L103
- Onodera, M., Renzini, A., Carollo, M., et al. 2012, arXiv:1206.1540
- Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., & Burkert, A. 2010, *ApJ*, 725, 23
- Oser, L., Naab, T., Ostriker, J. P., & Johansson, P. H. 2012, *ApJ*, 744, 63
- Patel, S. G., van Dokkum, P. G., Franx, M., et al. 2012, arXiv:1208.0341
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, 139, 2097
- Salim, S., Rich, R. M., Charlot, S., et al. 2007, *ApJS*, 173, 267
- Sánchez-Blázquez, P., et al. 2006, *MNRAS*, 371, 703
- Saracco, P., Gargiulo, A., & Longhetti, M. 2012, *MNRAS*, 422, 3107
- Sérsic, J. L. 1968, *Cordoba, Argentina: Observatorio Astronomico*, 1968,
- Shen, S., Mo, H. J., White, S. D. M., Blanton, M. R., Kauffmann, G., Voges, W., Brinchmann, J., & Csabai, I. 2003, *MNRAS*, 343, 978
- Skelton, R. et al. in preparation
- Szomoru, D., et al. 2010, *ApJ*, 714, L244
- Szomoru, D., Franx, M., & van Dokkum, P. G. 2012, *ApJ*, 749, 121
- Taylor, E. N., Franx, M., Glazebrook, K., et al. 2010a, *ApJ*, 720, 723
- Taylor, E. N., Franx, M., Brinchmann, J., van der Wel, A., & van Dokkum, P. G. 2010b, *ApJ*, 722, 1
- Toft, S., Gallazzi, A., Zirm, A., et al. 2012, *ApJ*, 754, 3
- Treu, T., Ellis, R. S., Liao, T. X., & van Dokkum, P. G. 2005a, *ApJ*, 622, L5
- Treu, T., Ellis, R. S., Liao, T. X., et al. 2005b, *ApJ*, 633, 174
- Trujillo, I., et al. 2006, *MNRAS*, 373, L36
- Trujillo, I., Cenarro, A. J., de Lorenzo-Cáceres, A., et al. 2009, *ApJ*, 692, L118
- van de Sande, J., Kriek, M., Franx, M., et al. 2011, *ApJ*, 736, L9
- van Dokkum, P. G. 2001, *PASP*, 113, 1420
- van Dokkum, P. G., et al. 2008, *ApJ*, 677, L5

- van Dokkum, P. G., Kriek, M., & Franx, M. 2009, *Nature*, 460, 717
- van Dokkum, P. G., et al. 2010, *ApJ*, 709, 1018
- van der Wel, A., Franx, M., van Dokkum, P. G., & Rix, H.-W. 2004, *ApJ*, 601, L5
- van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2005, *ApJ*, 631, 145
- van der Wel, A., Holden, B. P., Zirm, A. W., et al. 2008, *ApJ*, 688, 48
- van der Wel, A., Bell, E. F., van den Bosch, F. C., Gallazzi, A., & Rix, H.-W. 2009, *ApJ*, 698, 1232
- Warren, S. J., Hambly, N. C., Dye, S., et al. 2007, *MNRAS*, 375, 213
- Whitaker, K. E., et al. 2010, *ApJ*, 719, 1715
- Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, 735, 86
- Whitaker, K. E., Kriek, M., van Dokkum, P. G., et al. 2012, *ApJ*, 745, 179
- Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, 691, 1879
- Wuyts, S., van Dokkum, P. G., Kelson, D. D., Franx, M., & Illingworth, G. D. 2004, *ApJ*, 605, 677
- Wuyts, S., Labbé, I., Schreiber, N. M. F., et al. 2008, *ApJ*, 682, 985